FORM PTO-1390 (Rev. 11-2000) U.S. DEPARTMENT OF COMMERCE ATENT AND TRADEMARK OFFICE

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

ATTORNEY'S DOCKET NUMBER: 640100-430

U.S. APPLICATION NO. (If known, see 37 CFR 1.5): 09/937974

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/US00/08751	31 March 2000	1 April 1999

TITLE OF INVENTION: HUMAN MESENCHYMAL DNAs AND EXPRESSION PRODUCTS

APPLICANT(S) FOR DO/EO/US: Christian Van den Bos and Gabriel Mbalaviele

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☒ This express request to being national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (21) indicated below.
4. ☒ The US has been elected by the expiration of 19 months from the priority date (Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☐ is attached herewith (required only if not communicated by the International Bureau).
   b. ☐ has been communicated by the International Bureau.
   c. ☒ is not required, as the application was filed in the United States Receiving Office (RO/US)
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2))
   a. ☐ is attached hereto.
   b. ☐ has been previously submitted under 35 U.S.C. 154(d)(4).
7. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   b. ☐ have been communicated by the International Bureau.
   c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   d. ☐ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11 to 20 below concern other document(s) or information included:

11. ☒ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. ☒ A **FIRST** preliminary amendment.
14. ☐ A **SECOND or SUBSEQUENT** preliminary amendment.
15. ☐ A substitute specification.
16. ☐ A change of power of attorney and/or address letter.
17. ☒ A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C. 1.821-1.825.
18. ☒ A second copy of the published international application under 35 U.S.C. 154(d)(4).
19. ☐ A second copy of the English language translation of the international application under 35 U.S.C. 15(d)(4).
20. ☐ Other items or information:

U.S. APPLICATION NO. (If known, see 37 C.F.R. 1.5)	INTERNATIONAL APPLICATION NO.	ATTORNEY'S DOCKET NUMBER
09/937974	PCT/US00/08751	640100-430

21. ☒ The following fees are submitted:

		CALCULATIONS PTO USE ONLY
Basic National Fee (37 CFR 1.492(a)(1)-(5)): Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO and JPO	$1000.00	
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO	$860.00	
International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO	$710.00	
International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4)	$690.00	
International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4)	$100.00	

ENTER APPROPRIATE BASIC FEE AMOUNT =				$	1,040.00
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☒ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				$	130.00
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	$	
Total Claims	24 - 20 =	4	X $18.00	$	72.00
Independent Claims	6 - 3 =	3	X **$84.00**	$	252.00
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ **$270.00**	$	
TOTAL OF ABOVE CALCULATIONS			=	$	1364.00
☒ Applicant claims small entity status. See 37 CFR 1.27. The fees indicated above are reduced by 1/2				$	682.00
SUBTOTAL			=	$	682.00
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492(f))				$	
TOTAL NATIONAL FEE			=	$	682.00
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property				$	
TOTAL FEES ENCLOSED =				$	682.00
				Amount to be refunded:	$
				charged:	$

a. ☒ A check no. _____ in the amount of $682.00 to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. **03-0678** in the amount of $_____ to cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. **03-0678**. A duplicate copy of this sheet is enclosed.

d. ☐ Fees are to be charged to a credit card. **WARNING**: Information on this form may become public. **Credit Card information should not be included on this form.** Provide credit card information and authorization on PTO-2038.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Alan J. Grant, Esq.
Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart & Olstein
6 Becker Farm
Roseland, New Jersey 07068

Alan J. Grant
SIGNATURE

Alan J. Grant
NAME

33,389
REGISTRATION NUMBER

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
Patent Examining Operations

Applicant(s):	Van den Bos and Mbalaviele		
Serial No:	PCT/US00/08751	Art Unit:	Unassigned
Filed:	31 March 2000	Examiner:	Unassigned
Title:	HUMAN MESENCHYMAL DNAs AND EXPRESSION PRODUCTS		
Docket No:	640100-430		

Commissioner for Patents
Washington, D.C. 20231

Preliminary Amendment

Sir:

The above-referenced PCT application is filed herewith as a National-Stage application under 35 U.S.C. 371. Prior to examination on the merits, please enter the following amendment.

In the Specification:

Please amend the paragraph starting on page 1, line 8, to read as follows:

"This application is a national stage filing based on PCT/US00/08751 and claims the priority of U.S. Provisional Applications 60/148,800, filed 13 August 1999, and 60/127,418, filed 1 April 1999, the disclosures of which are hereby incorporated by reference in their entirety."

REMARKS

The application filed herewith is a national stage filing under 35 U.S.C. 371 and the specification has been amended to reflect this.

The Commissioner is authorized to charge payment of any additional filing fees required under 37 CFR 1.16 associated with this communication or credit any overpayment to Deposit Account No. 03-0678.

EXPRESS MAIL CERTIFICATE

Express Mail Label No. EL680647645US

Deposit Date: 1 October 2001

I hereby certify that this paper and the attachments hereto are being deposited today with the U.S. Postal Service "Express Mail Post Office To Addressee" service under 37 CFR 1.10 on the date indicated above addressed to:

Commissioner for Patents
Washington, DC 20231

Alan J. Grant — 10/1/01

Alan J. Grant, Esq. **Date**

Respectfully submitted,

Alan J. Grant

Alan J. Grant, Esq.
Reg. No. 33,389
CARELLA, BYRNE, BAIN, GILFILLAN,
CECCHI, STEWART & OLSTEIN
6 Becker Farm Road
Roseland, NJ 07068
Tel. No.: (973) 994-1700

AMENDED SPECIFICATION

The paragraph starting on page 1, line 8, has been amended as follows:

"This application is a national stage filing based on PCT/US00/08751 and claims the priority of U.S. Provisional Applications 60/148,800, filed 13 August 1999, and 60/127,418, filed 1 April 1999, the disclosures of which are hereby incorporated by reference in their entirety."

HUMAN MESENCHYMAL DNAs AND EXPRESSION PRODUCTS

This application claims the priority of U.S. Provisional Applications 60/148,800, filed 13 August 1999, and 60/127,418, filed 1 April 1999, the disclosures of which are hereby incorporated by reference in their entirety.

BACKGROUND OF THE INVENTION

This invention relates to newly identified polynucleotide sequences corresponding to transcription products of human genes, and to complete gene sequences associated therewith and to gene expression products thereof and to uses for the foregoing.

Osteoblasts, key cells in bone formation, or osteogenesis, are formed from mesenchymal stem cells. Such mesenchymal stem cells (or MSCs) of numerous mammalian species can be induced to differentiate into connective tissue cell lineages by varying the *in vitro* culture conditions. Osteogenesis, the differentiation into bone cells, has been reported as a means to generate replacement bone from cultured and implanted MSCs (Bruder et al, Growth Kinetics, Self-Renewal, and the Osteogenic Potential of Purified Human Mesenchymal Stem Cells During Extensive Subcultivation and Following Cryopreservation, J. Cell Biochem., 64(2):278-294 (Feb. 1997); Jaiswal et al., Osteogenic Differentiation of Purified, Culture-Expanded Human Mesenchymal Stem

Cells In Vitro, J. Cell Biochem., 64(2):295-312 (Feb. 1997), Kadiyala et al., Culture Expanded Canine Mesenchymal Stem Cells Possess Osteochondrogenic Potential In Vivo and In Vitro, Cell Transplant, 6(2):125-134 (Mar-Apr 1997)).

The process by which MSCs undergo osteogenic differentiation in culture is marked by the development of an osteoblastic morphology, the deposition of a hydroxyapetite mineralized extracellular matrix characteristic of osteoblasts and the presence of terminally differentiated osteocytes, as well as the expression of alkaline phosphatase (Jaiswal et al., Osteogenic Differentiation of Purified, Culture-Expanded Human Mesenchymal Stem Cells In Vitro, J. Cell Biochem., 64(2):295-312 (Feb. 1997)). Mechanisms underlying the osteogenic differentiation of human MSCs (hereafter, hMSCs) are poorly understood. Identification of proteins produced during this process would greatly facilitate the discovery and development of small molecules that target the osteoblast and its bone forming potential. Identification of these factors would be accelerated by the availability of relevant cDNA libraries constructed from hMSCs during various stages of their differentiation.

Identification and sequencing of human genes is a major goal of modern Molecular Biology. For example, by identifying genes and determining their sequences, scientists have been able to make large quantities of valuable human "gene products." These include human insulin, interferon, Factor VIII, tumor necrosis factor, human growth hormone, tissue plasminogen activator, and numerous other compounds. Additionally, knowledge of gene sequences can provide the key to treatment or cure of genetic diseases (such as muscular dystrophy and cystic fibrosis).

BRIEF SUMMARY OF THE INVENTION

In accordance with the present invention, Mesenchymal stem cells (MSCs) have been isolated and culture expanded from humans, and from them new cDNA libraries have been constructed from messenger ribonucleic acids (hereafter, mRNAs) isolated from hMSCs.

It is an object of the present invention to obtain cDNA libraries from purified and cultured MSCs and to use these isolated nucleic acids, isolated sequences, and fragments thereof, in the determination and preparation of the expression products of these nucleic acids and sequences, including fragments thereof.

It is a further object of the present invention to use the cDNAs so produced, and fragments thereof, as well as their expression products, as chromosomal markers for determining the location of genes within the genome, and alleles thereof, expressed during the development of differentiated mesenchymal cells.

It is yet another object of the present invention to provide DNA sequences for use in human "fingerprinting" whereby different individuals can be distinguished based on the sequences of the genes identified as wholly, or partly, identical to those disclosed herein.

It is still another object of the present invention to provide polynucleotide sequences corresponding to the genes coding for polypeptides as disclosed herein whereby such sequences can be compared with those found in similar chromosomal locations in animals, especially mammals, and most especially humans, where such animal is

afflicted with a disease affecting bone growth, or such other disease, or diseases, as may be affected by such genes, and thus detecting the presence of mutations in said genes leading to such diseases.

It is a still further object of the present invention to provide genetically engineered cells, and vectors, containing one or more copies of the nucleic acids, or DNAs, or genes, or nucleotide sequences according to the present invention, capable of expressing said peptides, or polypeptides, or proteins for rapid cloning of genes according to the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows the consensus sequence (SEQ ID NO: 27) for the novel DNA sequence of the invention as determined from different cDNA clones of said sequence, the latter being about 2.5 kb in length.

Figure 2 is a deduced amino acid sequence for the protein expressed from the sequence of Figure 1, residues 125 through 1717 and corresponding to SEQ ID NO:29. The amino acids set off between asterisks constitute a bipartite nuclear localization signal. The isoelectric point and molecular weight were also calculated for the putative protein.

Figure 3 shows the results of a dot blot assay for the presence of the novel DNA sequence in a variety of human tissues. For this assay, a prefabricated dot blot from Clontech (#7770-1) was hybridized using a probe generated from the 2.5 kb cDNA of Figure 1 and treated according to the manufacturer's instructions. Signals due to bound probe were analyzed using a Storm 860 phosphorimager and imagequant software.

Figure 4 is a bar graph showing the distribution of the sequence of Figure 1 in a variety of human tissues based on relative mRNA abundance. The highest signal strength was in cells of adult heart and lowest was in fetal thymus. The bar graphs were generated using data from the dot blots of Figure 3 and were imported into an Excel spreadsheet. The data were then analyzed as arbitrary signal strength per tissue after subtracting background (due to non-specific hybridization). The order of the tissues in the bar graph reflects signal strength (and therefor differs from that on the dot blot of Figure 3). Figure 4(b) is a continuation of Figure 4(a).

DETAILED DESCRIPTION OF THE INVENTION

One aspect of the present invention is directed to nucleic acids and isolated DNA sequences and molecules, and fragments thereof (and corresponding isolated RNA sequences, and fragments thereof), including sequences complementary to the foregoing, showing sequence similarity to, or capable of hybridizing to, the DNA sequences identified in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, and 27 or 28. The present invention is also directed to fragments or portions of such sequences which contain at least 15 bases, preferably at least 30 bases, more preferably at least 50 bases and most preferably at least 80 bases, and to those sequences which are at least 60%, preferably at least 80%, and most preferably at least 95%, especially 98%, identical thereto, and to DNA (or RNA) sequences encoding the polypeptides of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 29, including fragments and portions thereof and, when derived from natural sources, includes alleles thereof.

In accordance with the present invention, the term "percent identity" or "percent identical," when referring to a sequence, means that a sequence is compared to a claimed or described sequence after alignment of the sequence to be compared (the "Compared Sequence") with the described or claimed sequence (the "Reference Sequence"). The Percent Identity is then determined according to the following formula:

Percent Identity = 100 [1-(C/R)]

wherein C is the number of differences between the Reference Sequence and the Compared Sequence over the length of alignment between the Reference Sequence and the Compared Sequence wherein (i) each base or amino acid in the Reference Sequence that does not have a corresponding aligned base or amino acid in the Compared Sequence and (ii) each gap in the Reference Sequence and (iii) each aligned base or amino acid in the Reference Sequence that is different from an aligned base or amino acid in the Compared Sequence, constitutes a difference; and R is the number of bases or amino acids in the Reference Sequence over the length of the alignment with the Compared Sequence with any gap created in the Reference Sequence also being counted as a base or amino acid.

If an alignment exists between the Compared Sequence and the Reference Sequence in which the percent identity as calculated above is about equal to or greater than a specified minimum Percent Identity then the Compared Sequence has the specified minimum percent identity to the Reference Sequence even though alignments may exist in which the hereinabove calculated Percent Identity is less than the specified Percent Identity.

Yet another aspect of the present invention is directed to an isolated DNA (or RNA) sequence or molecule comprising at least the

coding region of a human gene (or a DNA sequence encoding the same polypeptide as such coding region), in particular an expressed human gene, which human gene comprises a DNA sequence homologous with, or contributing to, the sequence depicted in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25 and 27 or 28, or one at least 60%, preferably at least 80%, and most preferably at least 95%, especially 98%, identical thereto, including 100% identity, as well as fragments or portions of the coding region which encode a polypeptide having a similar function to the polypeptide encoded by said coding region. Thus, the isolated DNA (or RNA) sequence may include only the coding region of the expressed gene (or fragment or portion thereof as hereinabove indicated) or may further include all or a portion of the non-coding DNA (or RNA) of the expressed human gene.

In general, sequences homologous with and contributing to the sequences of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25 and 27 or 28 (or one at least 60%, preferably at least 80%, and most preferably at least 95% identical or homologous thereto) are from the coding region of a human gene.

The present invention also relates to vectors or plasmids which include such DNA (or RNA) sequences, as well as the use of the DNA (or RNA) sequences.

The sequences depicted in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25 and 28 are hybridizable with actual DNA and RNA sequences as derived from different human tissues. These sequences represent cDNA clones.

The sequence depicted in Figure 1 (SEQ ID NO: 27) is hybridizable with actual DNA and RNA sequences as derived from different human tissues. A number of cDNA clones have been generated. The nucleotide

sequence of Figure 1 (SEQ ID NO: 27) itself showed a nuclear location in the various tissues studied. The distribution of this sequence in various human tissues is shown in Figures 3 and 4. Some of these clones had an additional 3'-untranslated region, the presence of which is generally related to the extent to which the mRNA species remain in the cell before being turned over. See Kingman, Genetic Engineering, Blackwell, 1988, at page 313. The 3'-untranslated region may also regulate the frequency at which the mRNA is translated and thus constitute a mechanism by which the expression of the protein can be regulated. (Gray, N.K. & Wickens, M., Control of Translation Initiation in Animals, Ann. Rev. Cell Dev. Biol., 14:399-458 (1998).

The polynucleotides of the present invention may be in the form of RNA or in the form of DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand. The coding sequence which encodes the mature polypeptide may be identical to the coding sequences present as open reading frames (ORFs) of the spolynucleotide sequences disclosed herein or may be a different coding sequence, which coding sequence, as a result of the redundancy or degeneracy of the genetic code, encodes the same mature polypeptide as the polynucleotide sequences of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25 and 28.

The polynucleotides that code for the polypeptides disclosed herein as putative proteins SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26 and 29 may include, but are not limited to: only the coding sequence for the mature polypeptide; the coding sequence for the mature polypeptide and additional coding sequence such as a leader or secretory sequence, a proprotein sequence and a membrane anchor; the coding sequence for the mature polypeptide (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding

sequence 5' and/or 3' of the coding sequence for the mature polypeptide.

The polynucleotide which codes for the polypeptide of Figure 2 (SEQ ID NO:29) may include, but is not limited to: only the coding sequence for the mature polypeptide; the coding sequence for the mature polypeptide and additional coding sequence such as a leader or secretory sequence, a proprotein sequence and a membrane anchor; the coding sequence for the mature polypeptide (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding sequence 5' and/or 3' of the coding sequence for the mature polypeptide.

The term "polynucleotide" as used for the present invention encompasses a polynucleotide which includes only coding sequence for the polypeptide as well as a polynucleotide which includes additional coding and/or non-coding sequences.

The present invention further relates to variants of the hereinabove described polynucleotides which encode fragments, analogs and derivatives of the polypeptides having the amino acid sequences of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26 and 29. Variants of the polynucleotide may be naturally occurring allelic variants of the polynucleotides or a non-naturally occurring variant of the polynucleotides.

Thus, the nucleic acids, or polynucleotides, according to the present invention may have coding sequences which are naturally occurring allelic variants of the coding sequence shown in SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 and 28. As known in the art, an allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides, which does not substantially alter the function of the encoded polypeptide.

The present invention also includes polynucleotides, wherein the coding sequence for the mature polypeptide may be fused in the same reading frame to a polynucleotide sequence which aids in expression and secretion of a polypeptide from a host cell, for example, a leader sequence which functions as a secretory sequence for controlling transport of a polypeptide from the cell and a transmembrane anchor which facilitates attachment of the polypeptide to a cellular membrane. The polypeptide having a leader sequence is a preprotein and may have the leader sequence cleaved by the host cell to form the mature polypeptide. The polynucleotides may also encode for a proprotein which is the mature protein plus additional 5' amino acid residues. A mature protein having a prosequence is a proprotein and is often an inactive form of the protein. Once the prosequence is cleaved an active mature protein remains.

Thus, for example, the polynucleotide of the present invention may encode for a mature protein, for a protein having a prosequence, for a protein having a transmembrane anchor or for a polypeptide having a prosequence, a presequence (leader sequence) and a transmembrane anchor.

The polynucleotides of the present invention may also have the coding sequence fused in frame to a marker sequence which allows for purification of the polypeptide of the present invention. The marker sequence may be a hexa-histidine tag supplied by a pQE-9 vector to provide for purification of the mature polypeptide fused to the marker in the case of a bacterial host, or, for example, the marker sequence may be a hemagglutinin (HA) tag when a mammalian host, e.g. COS-7 cells, is used. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson, I., et al., Cell, 37:767 (1984)).

Fragments of the full length polynucleotide of the present invention may be used as hybridization probes for a cDNA library to isolate

the full length cDNA and to isolate other cDNAs which have a high sequence similarity to the gene or similar biological activity. Probes of this type preferably have at least 15 bases, may have at least 30 bases and even 50 or more bases. The probe may also be used to identify a cDNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete gene including regulatory and promotor regions, exons, and introns. An example of a screen comprises isolating the coding region of the gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention are used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

A polynucleotide according to the present invention may have at least 15 bases, preferably at least 30 bases, and more preferably at least 50 bases which hybridize to a polynucleotide of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 and 28 and which has an identity thereto, as hereinabove described, and which may or may not retain activity. Such polynucleotides may be employed as probes for the polynucleotides or genes coding for the polypeptides of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 29, for example, for recovery of the polynucleotide or as a diagnostic probe or as a PCR primer.

The polynucleotides according to the present invention may also occur in the form of mixtures of polynucleotides hybridizable to some extent with the gene sequences containing any of the nucleotide sequences of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 and 28 including any and all fragments thereof, and which polynucleotide mixtures may be composed of any number of such polynucleotides, or fragments thereof, including mixtures having at least 10, perhaps at least 30 such sequences, or fragments thereof.

Because coding regions comprise only a small portion of the human genome, identification and mapping of transcribed regions and coding regions of chromosomes is of significant interest. There is a corresponding need for reagents for identifying and marking coding regions and transcribed regions of chromosomes. Furthermore, such human sequences are valuable for chromosome mapping, human identification, identification of tissue type and origin, forensic identification, and locating disease-associated genes (i.e., genes that are associated with an inherited human disease, whether through mutation, deletion, or faulty gene expression) on the chromosome.

Various aspects of the present invention include each of the individual sequences, corresponding partial and complete cDNAs, genomic DNA, mRNA, antisense strands, PCR primers, coding regions, and constructs. Expression vectors and polypeptide expression products, are also within the scope of the present invention, along with antibodies, especially monoclonal antibodies, to such expression products.

As used herein and except as noted otherwise, all terms are defined as given below.

In accordance with the present invention, the term "gene" or "cistron" means the segment of DNA (or DNA segment) involved in producing a polypeptide chain; it includes regions preceding and following the coding region (5'-and 3'- untranslated regions, or UTRs, also called leader and trailer sequences, regions, or segments) as well as intervening sequences (introns) between individual coding segments (exons), which intronic regions are typically removed during processing of post-transcriptional RNA to form the final translatable mRNA product. Of course, by their nature, cDNAs contain no intronic sequences.

In accordance with the present invention, the term "DNA segment" refers to a DNA polymer, in the form of a separate fragment or as a component of a larger DNA construct, which has been derived from DNA isolated at least once in substantially pure form, i.e., free of contaminating endogenous materials and in a quantity or concentration enabling identification, manipulation, and recovery of the segment and its component nucleotide sequences by standard biochemical methods, for example, using a cloning vector. Such segments are provided in the form of an open reading frame uninterrupted by internal nontranslated sequences (introns), which are typically present in eukaryotic genes. Sequences of non-translated DNA may be present downstream from the open reading frame, where the same do not interfere with manipulation or expression of the coding regions.

The nucleic acids and polypeptide expression products disclosed according to the present invention, as well as expression vectors containing such nucleic acids, may be in "enriched form." As used herein, the term "enriched" means that the concentration of the material is at least about 2, 5, 10, 100, or 1000 times its natural concentration (for example), advantageously 0.01%, by weight, preferably at least about 0.1% by weight. Enriched preparations of about 0.5%, 1%, 5%, 10%, and 20% by weight are also contemplated. The sequences, constructs, vectors, clones, and other materials comprising the present invention can advantageously be in enriched or isolated form. For example, removal, via the differential display techniques described herein, of clones corresponding to ribosomal RNA and "housekeeping" genes and clones without human cDNA inserts results in a library that is "enriched" in the desired clones.

The DNA and RNA sequences, and polypeptides, disclosed in accordance with the present invention will commonly be in isolated form. The term "isolated" means that the material is removed from its original

environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide, or DNA, present in a living animal is not isolated, but the same polynucleotide or DNA, separated from some or all of the coexisting materials in the natural system, is isolated. Such DNA could be part of a vector and/or such polynucleotide could be part of a composition, and still be isolated in that such vector or polynucleotide is not part of its natural environment.

The DNA and RNA sequences, and polypeptides, disclosed in accordance with the present invention may also be in "purified" form. The term "purified" does not require absolute purity; rather, it is intended as a relative definition, and can include preparations that are highly purified or preparations that are only partially purified, as those terms are understood by those of skill in the relevant art. Individual clones isolated from a cDNA library have been conventionally purified to electrophoretic homogeneity. The cDNA clones are obtained via manipulation of a partially purified naturally occurring substance (messenger RNA). By conversion of mRNA into a cDNA library, pure individual cDNA clones can be isolated from the synthetic library by clonal selection. Thus, creating a cDNA library from RNA and subsequently isolating individual clones from that library results in an approximately 10^6 fold purification of the native message. Purification of starting material or natural material to at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated. Furthermore, claimed polynucleotide which has a purity of preferably 0.001%, or at least 0.01% or 0.1%; and even desirably 1% by weight or greater is expressly contemplated.

The term "coding region" refers to that portion of a human gene which either naturally or normally codes for the expression product of that gene in its natural genomic environment, i.e., the region coding *in vivo* for the native expression product of the gene. The coding region can

be from a normal, mutated or altered gene, or can even be from a DNA sequence, or gene, wholly synthesized in the laboratory using methods well known to those of skill in the art of DNA synthesis.

In accordance with the present invention, the term "nucleotide sequence" refers to a heteropolymer of deoxyribonucleotides. Generally, DNA segments encoding the proteins provided by this invention are assembled from cDNA fragments and short oligonucleotide linkers, or from a series of oligonucleotides, to provide a synthetic gene which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon.

The term "expression product" means that polypeptide or protein that is the natural transcription product of the gene and any nucleic acid sequence coding equivalents resulting from genetic code degeneracy and thus coding for the same amino acid(s).

The term "fragment" when referring to a coding sequence means a portion of DNA comprising less than the complete human coding region whose expression product retains essentially the same biological function or activity as the expression product of the complete coding region.

When referring to a portion of a polypeptide, as used herein, the terms "portion," "segment," and "fragment," refer to a continuous sequence of residues, such as amino acid residues, which sequence forms a subset of a larger sequence. For example, if a polypeptide were subjected to treatment with any of the common endopeptidases, such as trypsin or chymotrypsin, the oligopeptides resulting from such treatment would represent portions, segments or fragments of the starting polypeptide. Similarly, portions, segments or fragments of polynucleotides would include those products resulting from the treatment of such polynucleotides with

endonucleases.

The term "primer" means a short nucleic acid sequence that is paired with one strand of DNA and provides a free 3'OH end at which a DNA polymerase starts synthesis of a deoxyribonucleotide chain.

The term "promoter" means a region of DNA involved in binding of RNA polymerase to initiate transcription.

The term "open reading frame (ORF)" means a series of triplets coding for amino acids without any termination codons and is a sequence (potentially) translatable into protein.

The term "exon" means any segment of an interrupted gene that is represented in the mature RNA product.

As used herein, reference to a DNA sequence includes both single stranded and double stranded DNA. Thus, the specific sequence, unless the context indicates otherwise, refers to the single strand DNA of such sequence, the duplex of such sequence with its complement (double stranded DNA) and the complement of such sequence.

In accordance with the present invention, the overall approach to identification of cDNAs from hMSCs involved measurement of gene expression during growth of human mesenchymal stem cells in culture. Cells were harvested and the total RNA content thereof was recovered. Next, using various primer combinations, reverse transcriptase and polymerase chain reaction procedures (RT-PCR) were used to produce and amplify the corresponding cDNAs, which were then screened to find regulated DNA sequences that were subsequently purified and cloned. These clones were then sequenced and used to determine a consensus sequence (one based upon the most commonly occurring bases at each

nucleotide position in a sequence after the contributing sequences are aligned by residue position). The resulting sequences were then subjected to computer database searches for novelty, and any homology with known sequences, using, for example, the BLAST program and the GenBank database.

Using the RT-PCR methodology, the mRNA from the cells of interest (such as the hMSCs used in accordance with the present invention) is used to prepare a set or family of cDNAs corresponding to the expressed genes of the cell. This cDNA preparation is then exhaustively hybridized with mRNA of cells not expressing the gene and resulting in removal of all sequences from the cDNA preparation that are common to the two cell samples. All of the cDNA sequences that hybridize with the other mRNA and those that remain are then hybridized with mRNA from the cells expressing the gene (for example, cells from a healthy person or cells from tissues known to express the gene) to confirm that they are in fact the desired coding sequences. Because these latter clones contain sequences specific to the mRNA population of the cells of interest, they can subsequently be amplified and characterized using further rounds of PCR and the general techniques of molecular biology.

In accordance with the foregoing, a cDNA library was generated and corresponds to the sequences of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 and 28. Probes based on these cDNAs can be used to identify the relevant transcripts, using Northern Blotting Analysis methods well known in the art to localize these sequences within cells of various tissues. For example, the heaviest distribution of the gene coding for the polypeptide of Figure 2 (SEQ ID NO: 29) was in heart tissue, as shown in Figures 3 and 4.

In accordance with the present invention, cDNA was quantified by spotting 0.5 µl aliquots of standards and samples on ethidium agarose plates prepared as suggested in the instructions from the manufacturer (Stratagene, La Jolla, CA). Plates were incubated at room temperature for 15 minutes and DNA was visualized by UV transillumination. The respective cDNAs were then quantified by comparing spot intensities of the samples with those of the standards (the latter consisting of appropriate dilutions of 1 kb ladders (from Life Technology).

Aliquots of each amplified library were excised and plasmids from randomly chosen colonies were analyzed by restriction nuclease analysis. In accordance with the present invention, plasmid DNA was digested with both EcoRI and XhoI nucleases (New England Biolabs) and the resulting restriction fragments were separated on 1.5% agarose gel electrophoresis. The cDNA inserts ranged in size from less than 1 kbp to larger than 4 kbp (where 1 kbp = 1,000 nucleotide base pairs of duplex DNA).

Each of the DNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. The sequences can be used as diagnostic probes for the presence of a specific mRNA in a particular cell type as well as in genetic linkage analysis (polymorphisms). Further, the sequences can be used as probes for locating gene regions associated with genetic disease.

The nucleotide and gene sequences of the present invention are also valuable for chromosome identification. Each sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. The mapping of the

polynucleotides to specific chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-30 bp) from the sequences disclosed herein. Computer analysis of these sequences is used to rapidly select primers that do not span more than one exon in the corresponding genomic DNA, which would otherwise complicate the amplification process. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the sequences or subsequences disclosed herein will yield an amplified fragment.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more clones can be assigned per day using a single thermal cycler, as is well known in the art. Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map a sequence, or part of a sequence, to its chromosome include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes and preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence in situ hybridization (FISH) of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 500 or 600 bases; however, clones larger than 2,000 bp have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. FISH requires use of

the clone from which the sequence was derived, and the longer the better. For example, 2,000 bp is good, 4,000 is better, but more than 4,000 is probably not necessary to get good results a reasonable percentage of the time. For a review of this technique, see Verma et al., Human Chromosomes: a Manual of Basic Techniques. Pergamon Press, New York (1988).

Reagents for chromosome mapping can be used individually (to mark a single chromosome or a single site on that chromosome) or as panels of reagents (for marking multiple sites and/or multiple chromosomes). Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man (available on line through Johns Hopkins University Welch Medical Library)). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically close genes).

Next, it is necessary to determine if there are differences in the cDNA or genomic sequence between affected and unaffected individuals. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

With current resolution of physical mapping and genetic mapping techniques, a cDNA precisely localized to a chromosomal region

associated with the disease could be one of between 50 and 500 potential causative genes. (This assumes 1 megabase mapping resolution and one gene per 20 kb.)

Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that cDNA sequence. Ultimately, complete sequencing of genes from several individuals is required to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

In addition to the foregoing, the sequences of the invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide sequence to DNA or RNA. Polynucleotides suitable for use in these methods are usually 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al, Nucl. Acids Res., 6:3073 (1979); Cooney et al, Science, 241:456 (1988) ; and Dervan et al, Science, 251: 1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem., 56:560 (1991) ; Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix- formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide. Antisense RNA or oligonucleotide hybridization may also lead to RNAse H activation and hence destruction of the molecules involved in the hybrid.

The present invention is also a useful tool in gene therapy, which requires isolation of the disease-associated gene in question as a prerequisite to the insertion of a normal gene into an organism to correct a genetic defect. The high specificity of the cDNA probes according to this invention have promise of targeting such gene locations in a highly accurate manner.

The sequences of the present invention, as broadly defined, and including subsequences and fragments thereof, are also useful for identification of individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP.

However, RFLP is a pattern based technique, which does not require the DNA sequence of the individual to be sequenced. Portions of the sequences of the present invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. These sequences can also be used to prepare PCR primers for amplifying and isolating such selected DNA. One can, for example, take part of the sequence of the invention and prepare two PCR primers from the 5' and 3' ends of the sequence, or fragment of the sequence. These are used to amplify an individual's DNA, corresponding to the sequence. The amplified DNA is sequenced.

Panels of corresponding DNA sequences from individuals, made this way, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences, due to allelic differences. The sequences of the present invention can be used to particular advantage to obtain such identification sequences from individuals and from tissue. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the fragments or complete coding sequences comprising a part of the present invention can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals.

If a panel of reagents from the sequences according to the present invention is used to generate a unique ID database for an individual, those same reagents can later be used to identify tissue from that individual. Positive identification of that individual, living or dead can be made from extremely small tissue samples.

Another use for DNA-based identification techniques is in forensic biology. PCR technology can be used to amplify DNA sequences taken from very small biological samples. In one prior art technique, gene sequences are amplified at specific loci known to contain a large number of allelic variations, for example the DQα class II HLA gene (Erlich, H., PCR Technology, Freeman and Co. (1992)). Once this specific area of the genome is amplified, it is digested with one or more restriction enzymes to yield an identifying set of bands on a Southern blot probed with DNA corresponding to the DQα class II HLA gene. In accordance with the present invention, it is clear from the results depicted in Figure 3

and 4 that the novel gene signal according to the present invention is found in many different tissues of the body.

The sequences of the present invention can be used to provide polynucleotide reagents specifically targeted to additional loci in the human genome, and can enhance the reliability of DNA-based forensic identifications. Those sequences targeted to noncoding regions are particularly appropriate. As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Reagents for obtaining such sequence information are within the scope of the present invention. Such reagents can comprise complete genes, parts of genes or corresponding coding regions, or fragments of at least 15 bp, preferably at least 18 bp.

There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, in forensics when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar manner, these reagents can be used to screen tissue cultures for contamination.

Sequences that match perfectly to several different genes can be detected by hybridizing to chromosomes: if many chromosomal loci are observed, the sequence (or a close variant) is in more than one gene. This problem can be circumvented by using the 3'-untranslated part of the cDNA alone as a probe for the chromosomal location or for the full-length cDNA or gene. The 3'-untranslated region is more likely to be unique within gene families, since there is no evolutionary pressure to conserve a coding function of this region of the mRNA.

The cDNA libraries disclosed according to the present invention ideally use directional cloning methods so that either the 5' end of the cDNA (likely to contain coding sequence) or the 3' end (likely to be a non-coding sequence) can be selectively obtained.

Using the sequence information provided herein, the polynucleotides of the present invention can be derived from natural sources or synthesized using known methods. The sequences falling within the scope of the present invention are not limited to the specific sequences described, but include human allelic and species variations thereof. Allelic variations can be routinely determined by comparison of one sequence with a sequence from another individual of the same species. Furthermore, to accommodate codon variability, the invention includes sequences coding for the same amino acid sequences as do the specific sequences disclosed herein. In other words, in a coding region, substitution of one codon for another which encodes the same amino acid is expressly contemplated. (Coding regions can be determined through routine sequence analysis.)

In a cDNA library there are many species of mRNA represented. Each cDNA clone can be interesting in its own right, but must be isolated from the library before further experimentation can be completed. In order to sequence any specific cDNA, it must be removed and separated (i.e. isolated and purified) from all the other sequences. This can be accomplished by many techniques known to those of skill in the art. These procedures normally involve identification of a bacterial colony containing the cDNA of interest and further amplification of that bacteria. Once a cDNA is separated from the mixed clone library, it can be used as a template for further procedures such as nucleotide sequencing.

The present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

Thus, the present invention is not restricted to such constructs or sequences alone but also includes expression vehicles, which may include plasmids, viruses, or any other expression vectors, including cells and liposomes, containing any of the nucleic acids, nucleotide sequences, DNAs, RNAs, or fragments thereof, as disclosed according to the present invention. Furthermore, this will be true regardless of whether such sequences are coding sequences or non-coding sequences and whether such coding sequences code for all or part of the expression products as disclosed herein, so long as such expression products, or fragments thereof, exhibit some utility in keeping with the invention disclosed herein. Thus, while the present invention includes an isolated DNA sequence, or nucleic acid, that expresses a human protein when in a suitable expression system, for example, a cell-free, or *in vitro*, expression system, such system may also be contained in, or part of, a suitable expression vehicle, or vector, be that a cell, a plasmid, a virus, or other operative expression vector.

Such expression systems, especially where part of an expression vehicle, will commonly require some promoter region that may include a promoter different from that normally associated *in vivo* with the genes coding for the gene expression products and proteins disclosed according to the present invention. Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P_R, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

In a further embodiment, the present invention relates to host cells containing the above-described construct(s). The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a procaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, 1986)) .

The constructs in host cells can be used in a conventional manner to produce the gene product coded by the recombinant sequence. Alternatively, the encoded polypeptide, once the sequence is known from the cDNAs, or from isolation of the pure product, can be synthetically produced by conventional methods of peptide synthesis, either manual or automated.

Thus, in accordance with the present invention, once the coding sequence is known, or the gene is cloned which encodes the polypeptide, conventional techniques in molecular biology can be used to

obtain the polypeptide. More generally, the present invention includes all polypeptides coded for by any and each of the DNA or RNA sequences disclosed herein, including fragments of said polypeptides, as well as derivatives and functional analogs thereof.

At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. This is particularly useful in producing small peptides and fragments of larger polypeptides. (Fragments are useful, for example, in generating antibodies against the native polypeptide.)

Alternatively, the DNA encoding the desired polypeptide can be inserted into a host organism and expressed. The organism can be a bacterium, yeast, cell line, or multicellular plant or animal. The literature is replete with examples of suitable host organisms and expression techniques. For example, polynucleotide (DNA or mRNA) can be injected directly into muscle tissue of mammals, where it is expressed. This methodology can be used to deliver the polypeptide to the animal, or to generate an immune response against a foreign polypeptide. Wolff, et al., Science, 247:1465 (1990); Felgner, et al., Nature, 349:351 (1991). Alternatively, the coding sequence, together with appropriate regulatory regions (i.e., a construct), can be inserted into a vector, which is then used to transfect a cell. The cell (which may or may not be part of a larger organism) then expresses the polypeptide.

The present invention further relates to polypeptides having an amino acid sequence selected from SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 29, as well as fragments, analogs and derivatives of such polypeptide.

The terms "fragment," "derivative" and "analog," when referring to the polypeptides disclosed herein also mean polypeptides that

retain essentially the same biological function or activity as said polypeptides. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide. Such fragments, derivatives and analogs must have sufficient similarity to the polypeptides disclosed herein so that activity of the native polypeptide is retained.

The polypeptides of the present invention may be recombinant polypeptides, natural polypeptides or synthetic polypeptides, preferably recombinant polypeptides.

"Recombinant," as used herein, means that a protein is derived from recombinant (e.g., microbial or mammalian) expression systems. "Microbial" refers to recombinant proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Protein expressed in most bacterial cultures, e.g., E. coli, will be free of glycosylation modifications; protein expressed in yeast will have a glycosylation pattern different from that expressed in mammalian cells.

The fragment, derivative or analog of a polypeptide of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26 and 29 may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or

a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the abilities of those skilled in the art in view of the teachings herein.

The polypeptides of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity. When applied to polypeptides, the term "isolated" has its already stated meaning.

The polypeptides of the present invention include the polypeptides of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 29, in particular the mature polypeptide, as well as polypeptides which have at least 70% identity to these polypeptides, or which have, at least 90% identity to these polypeptides, still more preferably at least 95% identity to these polypeptides and also include portions of such polypeptides with such portion generally containing at least 30 amino acids and more preferably at least 50 amino acids.

Fragments or portions of the polypeptides of the present invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length polypeptides. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells are genetically engineered (transduced or

transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector, either of which may be in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the present invention. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The polynucleotides of the present invention may be employed for producing polypeptides by recombinant techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing a polypeptide. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

In accordance with the present invention, an appropriate DNA sequence or segment may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into the appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (for example, a promoter sequence) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the *E. coli. lac* or *trp*, the phage lambda P_L promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their

viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in E. coli.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as E. coli, Streptomyces, Salmonella typhimurium; fungal cells, such as yeast; insect cells such as Drosophila S2 and Spodoptera Sf9; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

"Recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. The expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use in yeast or eukaryotic expression systems preferably

include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an N-terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

"Recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extra chromosomally. The cells can be prokaryotic or eukaryotic. Recombinant expression systems as defined herein will express heterologous protein upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryatic and eukaryotic hosts are described by *Sambrook, et al., Molecular Cloning: A Laboratory Manual*, Second Edition, (Cold Spring Harbor, N.Y., 1989), Wu et al, *Methods in Gene Biotechnology* (CRC Press, New York, NY, 1997), and *Recombinant Gene Expression Protocols*, in *Methods in Molecular Biology*, Vol. 62, (Tuan, ed., Humana Press, Totowa, NJ, 1997), the disclosures of which are hereby incorporated by reference.

Transcription of the DNA encoding the polypeptides according to the present invention by higher eukarotes can be increased by insertion of an enhancer sequence into the vector. Such enhancers have been known for some time and are usually cis-acting elements of DNA, usually anywhere from 10 to 300 bp that act on a promoter to

increase transcription. Common examples include the SV40 enhancer, the cytomegalovirus early promoter enhancer, the polyoma enhancer and the enhancers found in adenovirus.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), α-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli, Bacillus subtilis, Salmonella typhimurium* and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

Recombinant protein produced in bacterial culture is conveniently isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

The protein, its fragments or other derivatives, or analogs thereof, or cells expressing them, can be used as an immunogen to produce antibodies thereto. These antibodies can be, for example, polyclonal, monoclonal, chimeric, single chain, Fab fragments, or the product of an Fab expression library. Various procedures known in the art may be used for the production of polyclonal antibodies.

Antibodies generated against the polypeptide corresponding to a sequence of the present invention can be obtained by direct injection of the polypeptide into an animal or by administering the polypeptide to an animal, preferably a nonhuman. The antibody so obtained will then bind the polypeptide itself. In this manner, even a sequence encoding only a fragment of the polypeptide can be used to generate antibodies binding the whole native polypeptide. Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide. Moreover, a panel of such antibodies, specific to a large number of polypeptides, can be used to identify and differentiate such tissue.

For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler and Milstein, 1975, Nature, 256:495-497), the trioma technique, the human B-cell

hybridoma technique (Kozbor et al., 1983, Immunology Today 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole, et al., 1985, in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention.

The antibodies can be used in methods relating to the localization and activity of the protein sequences of the invention, e.g., for imaging these proteins, measuring levels thereof in appropriate physiological samples and the like.

In carrying out the procedures of the present invention it is of course to be understood that reference to particular buffers, media, reagents, cells, culture conditions and the like are not intended to be limiting, but are to be read so as to include all related materials that one of ordinary skill in the art would recognize as being of interest or value in the particular context in which that discussion is presented. For example, it is often possible to substitute one buffer system or culture medium for another and still achieve similar, if not identical, results. Those of skill in the art will have sufficient knowledge of such systems and methodologies so as to be able, without undue experimentation, to make such substitutions as will optimally serve their purposes in using the methods and procedures disclosed herein.

Specific embodiments of the invention will now be further described in more detail in the following non-limiting examples and it will be appreciated that additional and different embodiments of the teachings of the present invention will doubtless suggest themselves to those of

skill in the art and such other embodiments are considered to have been inferred from the disclosure herein.

EXAMPLE

The proteins encoded by the nucleotide sequences of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 and 28 are expressed in U2OS cells. This is achieved by selectively PCR amplifying the coding regions thereof (based on the available open reading frames) and then cloning the resulting amplicon into a suitable mammalian expression vector. One such vector is pcDNA3.1 (sold by Invitrogen - #K4800-01). The expression of the protein encoded by the described polynucleotide sequence is detected in either of two ways: by use of specific antibodies raised against peptides derived from the amino acid sequence or by use of antibodies against tags added during the cloning procedure. Examples of such tags are the V5 epitope or a poly-histidine sequence as contained in the pcDNA3.1 vector. In order to accomplish this, cells will normally be transfected with the expression construct and cultured for 1 to 5 days. Cells will then be lysed and their protein content analyzed by western blotting using the above antibodies as appropriate. Cells will also be analyzed for the subcellular localization of the protein encoded by the described polynucleotide sequence by transfecting cells in suitable chambers, culturing them for 1 to 5 days and fixing them *in situ*. Such cells will then be analyzed for the presence and localization of the encoded protein by staining cells with the above-referenced antibodies. Alternatively, cells will be transfected with an expression system in which the protein encoded by the described polynucleotide sequence is fused to a directly detectable tag such as green fluorescent protein (GFP). The expression and localization of the protein encoded by the described polynucleotide sequence is then detected by analyzing that of GFP.

For purposes of identification of the polypeptides disclosed herein, each such polypeptide is listed in the table below along with its calculated molecular weight (Daltons) and its expected isoelectric point (pI).

Table 1.

SEQ ID NO:	# Residues	Mol. Wt.	pI
2	410	45786.9	8.96
4	227	26152.3	8.48
6	275	30781.6	10.00
10	84	8913.2	9.35
12	281	30386.7	9.35
14	322	32977.3	9.27
16	141	16444.4	9.34
18	219	24418.4	9.07
22	56	6356.3	7.85
24	344	37375.6	5.82
26	208	23864.9	9.71
29	531	60,576.6	9.63

The polypeptides of SEQ ID NOS: 8 and 20 corresponded only to partial sequences and thus no values could be calculated and such sequences are not in the table.

All of the polynucleotides from which these polypeptide sequences are derived are cDNAs isolated during a differential screen of osteogenic mesenchymal stem cells (MSCs) cultured for 4 days in the presence of osteogenic supplements.

SEQUENCE LISTING

<110> Van Den Boss, Christian
Mbalaviele, Gabriel

<120> Human Mesenchymal DNAs and Expression Products

<130> 640100-363

<140>
<141>

<150> 60/127,418
<151> 1999-04-01

<150> 60/148,800
<151> 1999-08-13

<160> 29

<170> PatentIn Ver. 2.1

<210> 1
<211> 1974
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived
from human mesenchymal stem cells after treatment
with growth supplement

<400> 1
ggcacgagca gaacaagttt ggcctgtcac tgctcctcat cctcctgagc cgtggtgaag 60
acctacagag ttcagaccct gctacagaat caacacaaaa taatcagtgg acggaggtga 120
tgttcatggc aacacgagaa cttctgcgga ttccccaagc agccctggcc aagccaatct 180
ctatacctac aaacctagtg tccctctttt ctcgctatgt tgaccggcag aaactgaact 240
tgctggagac aaaactgcag ctagttcagg ggatacgata aaagatctcc aaatgtgtcc 300
tgtacctcct tttggctgcc acctgcactg ctgccatcac caatggagtg tttttaatga 360
gggaaggaag gtagcttttt ccccaaagca aagtcttgtg ggatcgattc ctgtttacag 420
gggttgtctc tctaaatgtc agatatttcc ccactgctct atgaaatttg gctgggtgat 480
acttctgctg gtttctttac cttctgtgtt acagttctgc atgtcctact tttactcagt 540
tctgttttgc attttctttg ccctagagac acaagtgtaa tctctccctt tatccctcca 600
ctactccacc tcagagtaga ttgtagcctg ccaaaggatt ccttccctca tcctattgaa 660
gttgtttttt cattgcccca tattaatatg actatagaag agccaattaa gtagaatcaa 720
gatatacaca cacacataga tacacacaca cacaccccat acatgtattt atgtggtctt 780
cagagggtcc ttaaagaatg aattttagat tgaaaaatat ttagttgtct cattacctct 840
tctaaacaca aaccagctga tgtattttaa tctgtttctg ttctatcttg taattaattt 900
ggtgggttct acttgtttta acataaataa agagtatgca gcacgtttaa taaaatcaga 960
actcttaatt ggcttatgcc caggtctagg ctgagaagtc ctttttcttc ttcccacctt 1020
tatttcctta gtttctgtcc accttaatcg aaacaacaca tggttatgtc tttttcctgc 1080
tacaactaca gggtacttga gcctttcccc tcaagtgcat tcgaagtcac ccaggatgat 1140
cctcactagt agcctgcttt ggcagtgtgg ctttttgcac acttgccctg tcttcctgag 1200
actacttcag taagccatgc ttccttcttc cccactttta tttggtgtca tgaatagaaa 1260
cttccaaatg taaccatgga agctaagttt ggcctgcttt gctttttagt ctccacacca 1320
tgggcagaac tgctgtcttt actacttcat ctcacccaag tcccgttccc aggcagccag 1380
gggcctgggt ttgaataatt gcagggccag cctgccatga tctttctcac ttactcctct 1440
cccattcagc aatcaaccag actaaggagt tttgatccct agtgattaca gccctgaaga 1500
aaattaaatc tgaattaatt ttacatggcc ttcgtgatct ttctgctgtt cttacttttt 1560
cgaatgtagt tgggggtgtg gagggacagg ttatggtatt taaagagaat aaacattttg 1620
cacatacatg tattgtacaa cagtaagatc ctctgttaaa accagctgtc ctgttctcca 1680

```
tctccatttc ttcccatgct gtaaccccag gctccaccag ctgttcccca gtgatgttac 1740
ctagcttccc tctaccgttg tctactgacc atttccacta catgcctttc ctaccttccc 1800
ttcacaacca atcaagtgaa tacttgatta ttatctcttc cttactgtgc tttatctttt 1860
ttgtttggat tggttctaat taatgaaaat aaaagtttct aaatttacat ttttataggg 1920
tattgtaaat aaaaacaaat tgtatactta aaaaaaaaaa aaaaaaaaaa aaaa       1974
```

<210> 2
<211> 410
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative protein derived from ORF cDNA of SEQ ID NO:1

<400> 2

```
Met Glu Leu Pro Ser Gly Pro Gly Pro Glu Arg Leu Phe Asp Ser His
  1               5                  10                  15

Arg Leu Pro Gly Asp Cys Phe Leu Leu Leu Val Leu Leu Leu Tyr Ala
             20                  25                  30

Pro Val Gly Phe Cys Leu Leu Val Leu Arg Leu Phe Leu Gly Ile His
         35                  40                  45

Val Phe Leu Val Ser Cys Ala Leu Pro Asp Ser Val Leu Arg Arg Phe
     50                  55                  60

Val Val Arg Thr Met Cys Ala Val Leu Gly Leu Val Ala Arg Gln Glu
 65                  70                  75                  80

Asp Ser Gly Leu Arg Asp His Ser Val Arg Val Leu Ile Ser Asn His
                 85                  90                  95

Val Thr Pro Phe Asp His Asn Ile Val Asn Leu Leu Thr Thr Cys Ser
            100                 105                 110

Thr Pro Leu Leu Asn Ser Pro Pro Ser Phe Val Cys Trp Ser Arg Gly
        115                 120                 125

Phe Met Glu Met Asn Gly Arg Gly Glu Leu Val Glu Ser Leu Lys Arg
    130                 135                 140

Phe Cys Ala Ser Thr Arg Leu Pro Pro Thr Pro Leu Leu Leu Phe Pro
145                 150                 155                 160

Glu Glu Glu Ala Thr Asn Gly Arg Glu Gly Leu Leu Arg Phe Ser Ser
                165                 170                 175

Trp Pro Phe Ser Ile Gln Asp Val Val Gln Pro Leu Thr Leu Gln Val
            180                 185                 190

Gln Arg Pro Leu Val Ser Val Thr Val Ser Asp Ala Ser Trp Val Ser
        195                 200                 205

Glu Leu Leu Trp Ser Leu Phe Val Pro Phe Thr Val Tyr Gln Val Arg
    210                 215                 220

Trp Leu Arg Pro Val His Arg Gln Leu Gly Glu Ala Asn Glu Glu Phe
225                 230                 235                 240
```

```
Ala Leu Arg Val Gln Gln Leu Val Ala Lys Glu Leu Gly Gln Thr Gly
                245                 250                 255

Thr Arg Leu Thr Pro Ala Asp Lys Ala Glu His Met Lys Arg Gln Arg
            260                 265                 270

His Pro Arg Leu Arg Pro Gln Ser Ala Gln Ser Ser Phe Pro Pro Ser
        275                 280                 285

Pro Gly Pro Ser Pro Asp Val Gln Leu Ala Thr Leu Ala Gln Arg Val
    290                 295                 300

Lys Glu Val Leu Pro His Val Pro Leu Gly Val Ile Gln Arg Asp Leu
305                 310                 315                 320

Ala Lys Thr Gly Cys Val Asp Leu Thr Ile Thr Asn Leu Leu Glu Gly
                325                 330                 335

Ala Val Ala Phe Met Pro Glu Asp Ile Thr Lys Gly Thr Gln Ser Leu
            340                 345                 350

Pro Thr Ala Ser Ala Ser Lys Phe Pro Ser Ser Gly Pro Val Thr Pro
        355                 360                 365

Gln Pro Thr Ala Leu Thr Phe Ala Lys Ser Ser Trp Ala Arg Gln Glu
    370                 375                 380

Ser Leu Gln Glu Arg Lys Gln Ala Leu Tyr Glu Tyr Ala Arg Arg Arg
385                 390                 395                 400

Phe Thr Glu Arg Arg Ala Gln Glu Ala Asp
                405                 410
```

<210> 3
<211> 1619
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived
from human mesenchymal stem cells after treatment
with growth supplement

<400> 3

```
ggcacgaggc gccttctcct tgcttctggg ggtcgtggcc ttgctcccgc tgtgcgggaa 60
aagaatccag gcccttccac gcgcgtgtgg gtgcgggggc cccgaagtgc tcgtggttcc 120
ccgctaggtc tccgctgggg caggaaccgg aatcatgggt gggaccacca gcacccgccg 180
ggtcaccttc gaggcggacg agaatgagaa catcaccgtg gtgaagggca tccggctttc 240
ggaaaatgtg attgatcgaa tgaaggaatc ctctccatct ggttcgaagt ctcagcggta 300
ttctggtgct tatggtgcct cagtttctga tgaagaattg aaaagaagag tagctgagga 360
gctggcattg gagcaagcca agaaagaatc cgaagatcag aaacgactaa agcaagccaa 420
agagctggac cgagagaggg ctgctgccaa tgagcagtta accagagcca tccttcggga 480
gaggatatgt agcgaggagg aacgcgctaa ggcaaagcac ctggctaggc agctggaaga 540
gaaagaccga gtgctaaaga agcaggatgc attctacaaa gaacagctgg ctagactgga 600
ggagaggagc tcagagttct acagagtcac cactgaacaa tatcagaaag ctgctgaaga 660
gtggaagcaa agttcaagcg atatgagtct catccagtct gtgctgatct gcaggccaaa 720
attcttcagt gttaccgtga gaacacccac cagaccctca aatgctccgc tctggccacc 780
cagtatatgc actgtgtcaa tcatgccaaa cagagcatgc ttgagaaggg aggataaaaa 840
ctttcagaat gagcaaaaca ccatcaacgt taattccaga gatggaacat ttttttttcct 900
```

agtgagaaaa caacccattt gaagagaaga ccactaatga gaagaccact aaagagagac 960
atcaagaatg gattcagcag aatcatttca cgttttgaac agcagcagtt tgaagggcca 1020
aagccttgat cagggatcag tcattaaagg acactcttga gtattagtaa accctcttat 1080
gatgattaaa agagaagggc agcccctctcc acctttttggt actttctatt caacttgcac 1140
tgaccataaa atgtttctct tctgaacaag ccccatcatt tggtgaacct ccaccctaac 1200
aaagtaggat ggggttgggg gctaaattaa ttggagtggg gcgaggagag agccagaaaa 1260
catagatccg agggcagcag tgctgggtgg agagagccag aaaacagatc tggaggcagc 1320
agtgctggat ggaattgtct aggctgtggc atgttggttt tgtctttctt ttctcctttg 1380
attatgtaag agctatttca ttataactta ttatggtgat tatacaggca agaagacaaa 1440
aaggagagaa aatgtacctc ttctactgga ataatgttta tgattacaag tgagataagg 1500
tatttttatc aatatgaagg caaccttggc tgataaaacc tctatagtga atactcacat 1560
ctttacttca ctcactatca ataataaata tattttctga caaaaaaaaa aaaaaaaaa 1619

<210> 4
<211> 227
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative
protein derived from ORF of cDNA of SEQ ID NO: 3.

<400> 4
Met Gly Gly Thr Thr Ser Thr Arg Arg Val Thr Phe Glu Ala Asp Glu
1 5 10 15

Asn Glu Asn Ile Thr Val Val Lys Gly Ile Arg Leu Ser Glu Asn Val
20 25 30

Ile Asp Arg Met Lys Glu Ser Ser Pro Ser Gly Ser Lys Ser Gln Arg
35 40 45

Tyr Ser Gly Ala Tyr Gly Ala Ser Val Ser Asp Glu Glu Leu Lys Arg
50 55 60

Arg Val Ala Glu Glu Leu Ala Leu Glu Gln Ala Lys Lys Glu Ser Glu
65 70 75 80

Asp Gln Lys Arg Leu Lys Gln Ala Lys Glu Leu Asp Arg Glu Arg Ala
85 90 95

Ala Ala Asn Glu Gln Leu Thr Arg Ala Ile Leu Arg Glu Arg Ile Cys
100 105 110

Ser Glu Glu Glu Arg Ala Lys Ala Lys His Leu Ala Arg Gln Leu Glu
115 120 125

Glu Lys Asp Arg Val Leu Lys Lys Gln Asp Ala Phe Tyr Lys Glu Gln
130 135 140

Leu Ala Arg Leu Glu Glu Arg Ser Ser Glu Phe Tyr Arg Val Thr Thr
145 150 155 160

Glu Gln Tyr Gln Lys Ala Ala Glu Glu Val Glu Ala Lys Phe Lys Arg
165 170 175

Tyr Glu Ser His Pro Val Cys Ala Asp Leu Gln Ala Lys Ile Leu Gln
180 185 190

Cys Tyr Arg Glu Asn Thr His Gln Thr Leu Lys Cys Ser Ala Leu Ala

195 200 205

Thr Gln Tyr Met His Cys Val Asn His Ala Lys Gln Ser Met Leu Glu
210 215 220

Lys Gly Gly
225

<210> 5
<211> 1917
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived from human mesenchymal stem cells after treatment with growth supplement

<400> 5
ggcacgagga gagaggtggt gaaactcaag atccccatgg acaacaagga ggtcccgggc 60
gaggcgcccg cgccgtccgc cgacccggcg cgtccccacg cgtgccccga ctgcggccgc 120
gccttcgcgc gccgctccac gctggcgaag cacgcgcgca cgcacacggg cgaacggccc 180
ttcgggtgca ccgagtgcgg gcggcgcttc tcacagaagt cggcgctgac caaacacggc 240
cgcacgcaca cgggcgagcg gccctacgag tgccccgagt gcgacaaacg cttctcggcc 300
gcctcgaacc tgcggcagca ccgacggcgg cacacgggcg agaagccgta cgcatgcgcg 360
cactgcggcc gccgcttcgc gcagagctcc aactacgcac agcacctgcg cgtgcacacg 420
ggcgagaagc cgtacgcgtg cccggactgc ggacgcgcct ttggcggcag ctcgtgcctg 480
gcgcgccacc gacgcacgca cacgggcgag cggccctacg cttgcgccga ctgcggcacg 540
cgcttcgctc agagctcggc gctggccaag caccggcgcg tgcacacggg cgagaagccg 600
caccgctgcg ctgtgtgtgg ccgtcgcttc ggccaccgct ccaacctggc ggagcacgcg 660
cgcacgcaca caggcgagcg gccctacccc tgcgccgagt gcggccgccg cttccgccta 720
agctcgcact tcattcgcca ccgacgcgcg cacatgcggc gccgcctgta tatttgcgcc 780
ggctgcggca gggacttcaa gctgcccccc ggcgccacgg ccgccactgc caccgagcgt 840
tgcccggagt gtgagggcag ctgagtcccg cagggctgcg gaggggcgcg ctggggcttc 900
gacctggctg cactaaccca ggctcctcct cgccccggcc tccgggtctg ggaaattgag 960
gggacggcag gcccggctgc cctggaactg ggagacaggg agaatccccc gccggggtcc 1020
ctggaaacag tgcccacccc acatcactac attccctcgg cccgtgttag tgaataaagt 1080
attatatcct caccccaccc gtgcctgtga gtgaggtggg tgggagagga agaaagttgg 1140
ggttctccag gctcaggtgc caagtgagtt gtcaaggaac caaatgggga tgtaaaccta 1200
aaaggggttc ccggcacctc ggtttgtgtt ggttggaggt gatcgcacac ttggcccttg 1260
gttacgtcct cataacctta gacctgaaag ggcccataaa tatactatgt tcacgatcag 1320
acacgcactg cattcggcag agctccagtg agcaaggcac gaccctcaga tctcagtcta 1380
gtgaaggaga gaaaactgta ataacactac gttaaaggtt ttaactgctt tgttatgtaa 1440
gcttacccag cccggcgcac agtgactcac gcctgtaatc ccagcacttt gggagggcga 1500
ggctagcaga tcacttgagg ttaggagttc gataccagcc tggccaacat ggtgaaaccc 1560
ggtctctact aaaaatacaa aaattaactg ggtgtggtgg cgggcgcctg taatcccagc 1620
tactgagggg gctgaggcat gagaatcact tgaacctggg agacagaggt tgcaatgaac 1680
cgagatagtg ccattgcact ccggcctggg caacagagga agactgcctc aaacaaacaa 1740
aaaacaacaa accaaaccaa accaaaaaaa tctcaaagcg attggaccta gcagctcatg 1800
cctgtaatct ccagcacttt gggaggcgga ggcaggagga tctcttgaag tcaagagttt 1860
gagatcagcc tggagaacaa agtgagaccc ccatctatta aaaaaaaaaa aaaaaaa 1917

<210> 6
<211> 275
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative

protein derived from ORF of cDNA of SEQ ID NO: 5.

<400> 6

```
Met Asp Asn Lys Glu Val Pro Gly Glu Ala Pro Ala Pro Ser Ala Asp
 1               5                  10                  15

Pro Ala Arg Pro His Ala Cys Pro Asp Cys Gly Arg Ala Phe Ala Arg
            20                  25                  30

Arg Ser Thr Leu Ala Lys His Ala Arg Thr His Thr Gly Glu Arg Pro
        35                  40                  45

Phe Gly Cys Thr Glu Cys Gly Arg Arg Phe Ser Gln Lys Ser Ala Leu
    50                  55                  60

Thr Lys His Gly Arg Thr His Thr Gly Glu Arg Pro Tyr Glu Cys Pro
65                  70                  75                  80

Glu Cys Asp Lys Arg Phe Ser Ala Ala Ser Asn Leu Arg Gln His Arg
                85                  90                  95

Arg Arg His Thr Gly Glu Lys Pro Tyr Ala Cys Ala His Cys Gly Arg
            100                 105                 110

Arg Phe Ala Gln Ser Ser Asn Tyr Ala Gln His Leu Arg Val His Thr
        115                 120                 125

Gly Glu Lys Pro Tyr Ala Cys Pro Asp Cys Gly Arg Ala Phe Gly Gly
    130                 135                 140

Ser Ser Cys Leu Ala Arg His Arg Arg Thr His Thr Gly Glu Arg Pro
145                 150                 155                 160

Tyr Ala Cys Ala Asp Cys Gly Thr Arg Phe Ala Gln Ser Ser Ala Leu
                165                 170                 175

Ala Lys His Arg Arg Val His Thr Gly Glu Lys Pro His Arg Cys Ala
            180                 185                 190

Val Cys Gly Arg Arg Phe Gly His Arg Ser Asn Leu Ala Glu His Ala
        195                 200                 205

Arg Thr His Thr Gly Glu Arg Pro Tyr Pro Cys Ala Glu Cys Gly Arg
    210                 215                 220

Arg Phe Arg Leu Ser Ser His Phe Ile Arg His Arg Arg Ala His Met
225                 230                 235                 240

Arg Arg Arg Leu Tyr Ile Cys Ala Gly Cys Gly Arg Asp Phe Lys Leu
                245                 250                 255

Pro Pro Gly Ala Thr Ala Ala Thr Ala Thr Glu Arg Cys Pro Glu Cys
            260                 265                 270

Glu Gly Ser
        275
```

<210> 7
<211> 1150
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived
from human mesenchymal stem cells after treatment
with growth supplement

<400> 7
ggcacgagct gcggaaaatc tttccgctca gcacatcacc tcaaagtgca tctgcgagtg 60
cacacaggcg agcggcccta caagtgtccg cactgcgact acgcgggcac ccagtccggc 120
tcgctcaagt atcacctaca gcgccaccac cgggagcaga ggagcggggc cggccccggg 180
ccacccccgg agccaccgcc tccttcccag cggggttcgg ccccgcaatc tggagccaag 240
ccgtctccgc agcctgcgac ctgggtggag ggcgcctcaa gtccccggcc tccttctagc 300
ggtgctgggc cggggtcccg tcggaagccc gccagccctg ggaggaccct gcgcaacggg 360
cgaggcggtg aggccgaacc cctggacctg tccttgcggg cagggccggg aggcgaggcc 420
gggcctgggg gtgccctcca ccgctgcctc ttctgcccgt tcgccactgg agccccagag 480
ctcatggcct tgcaccttca agtgcaccac agccgccggg ctaggggccg ccggccaccc 540
caggctgacg cgtccccgcc ctatgcccga gtaccatcag gagagacccc tcccagtcct 600
tcgcaggaag gggaggaggg ctccgggctg tccagacccg gagaggcagg gctgggggg 660
caagaacggt agtgggccct caggggcgat tagcttagtg agcttacccc gccggagcgg 720
gggagcgcta gaggtagttg ggtagagaag ccagcagggg gcagcgtggg acccatatcc 780
cagccccagg cggaccagag gtggaaggcg cggcgggcgt aggagggtgg gctggcggta 840
cccacccagc ccaagggagt cacagtgcct tagaagtggc agaggtaggg tcaaagaacg 900
gggtccctgg gctggcagag aaggtttgtg tccctgttga ggagccactc tgccagtctt 960
tgctggtcca taggcgtgag agtttatttt tgtacaaggg gtgagggtgg ggggcactgt 1020
acccttctag ccccatcagg ggccctgtag acgtcgctat ttttggtacg attcctgtca 1080
tccctattgc agagtcgtgt ccccaataaa caggtgtctt gaggcacagg aaaaaaaaaa 1140
aaaaaaaaaa 1150

<210> 8
<211> 100
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative
protein derived from ORF of cDNA of SEQ ID NO: 7.

<400> 8
Ile Arg His Glu Leu Arg Lys Ile Phe Pro Leu Ser Thr Ser Pro Gln
1 5 10 15

Ser Ala Ser Ala Ser Ala His Arg Arg Ala Ala Leu Gln Val Ser Ala
20 25 30

Leu Arg Leu Arg Gly His Pro Val Arg Leu Ala Gln Val Ser Pro Thr
35 40 45

Ala Pro Pro Pro Gly Ala Glu Glu Arg Gly Arg Pro Arg Ala Thr Pro
50 55 60

Gly Ala Thr Ala Ser Phe Pro Ala Gly Phe Gly Pro Ala Ile Trp Ser
65 70 75 80

Gln Ala Val Ser Ala Ala Cys Asp Leu Gly Gly Gly Arg Leu Lys Ser
85 90 95

Pro Ala Ser Phe
100

<210> 9
<211> 1573
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived
from human mesenchymal stem cells after treatment
with growth supplement

<400> 9
ggcacgaggt gggctgcacg ctcccctcca tccccatcgg ccctgtcccc tggagtgtgt 60
cagagcccag gggagaatgc agcccaccag gagcacctgg acccccctgcc cgccacatgg 120
tgtggccatc actcagcccc taccccctgcc ctgctcctaa gggtagaaaa ctccagggtc 180
ccctgccacc gactgcccag ccactccaag ccccctggca gctgcccctc ctggagcaga 240
aagtgccttt atctcagcca tccgcagact gctcggccag atgcggggac aggctggaat 300
gagggaggcg tcttcatctc cctgccatcc ccctctcacg ccacccccgc ccccaccggg 360
ctgcaggtgc tgctgatgcg ctgggatctg attgaggata aaaaggaagg agagatgacc 420
cctacccccct catcccccag ttttgaaaag gtctaagcaa gtgagtctgg tggaggagct 480
gagggaggga gccatggaag gtgccagaag gaaggttggc gggggcacgt gtgggccgtg 540
gcttgggctg gtcagagtgg cgtgagctgc ccggcgcctg ccctgcccaa gtgaccaggg 600
aagtgtgtgt gtgtccatgt gtatgcgtgt ccgtctgtct gtctagtgtc tgggtttggc 660
ccaagactgg gctgtagtta cattaatgcc cagccagcca cccctgccac tcacccctcc 720
tggcccaggc cttgctgact ctctgagctg gggaggtggg aggccaggcg agcctgactc 780
tgttgatcta cccgtgcctg ggcccctccc ctcagagccc atggtaacga acccctagaa 840
aggagagaac gggcgtcagg ggtgcacagt ccacagctga agagcaaggt ttcgtggcag 900
cacggcccgg cccctcaccc tctgtcccca cgaggggacc catgggggct gtctttgcag 960
ggcacagatg accaaagtcc cttcctgctt cctgttacct gtcttgctcc tggggagaaa 1020
gaggggcctg atgagactcc actcaggtgc acacatcacc aggtgcatct gcaggcactg 1080
ggctggctgc ttgcagccag gagaaggtca gcgagaagga gtgtatgagt gtgagtgtgt 1140
gtgcatggaa gttggggcac tgggcgtctg actccctccc cacccaagag aggaaggacc 1200
cctcaccacc cccactggcg agacagttta ctttgccgac ttgccatgtt tttgccaaaa 1260
ccaagatttt gaaggaaatg agtggccagc gccagggccc aggccatgtg gcctgcccag 1320
cctcaatgtc acttggtggc ggggtggggt gggggtgggc agcagcatcc cagccttgag 1380
atgcttcact ttccttctct gtaaccagac tttgaaaaat tgttcgtttc atcaggctct 1440
gttcctcaat ggccttttgc tacgtgcctc ccgagaaatt tgtcttttg tataaatgac 1500
aaagtgttga aaatgtattt cctgaaataa atgtttcaaa tgcagaaacc caaaaaaaaa 1560
aaaaaaaaaa aaa 1573

<210> 10
<211> 84
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative
protein derived from ORF of cDNA of SEQ ID NO: 9.

<400> 10
Met Ser Val Ser Val Cys Ala Trp Lys Leu Gly His Trp Ala Ser Asp
1 5 10 15

Ser Leu Pro Thr Gln Glu Arg Lys Asp Pro Ser Pro Pro Pro Leu Ala
20 25 30

Arg Gln Phe Thr Leu Pro Thr Cys His Val Phe Ala Lys Thr Lys Ile
35 40 45

Leu Lys Glu Met Ser Gly Gln Arg Gln Gly Pro Gly His Val Ala Cys

50 55 60

Pro Ala Ser Met Ser Leu Gly Gly Gly Val Gly Trp Gly Trp Ala Ala
65 70 75 80

Ala Ser Gln Pro

<210> 11
<211> 1808
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived from human mesenchymal stem cells after treatment with growth supplement

<400> 11
ggcacgaggc cgcccggccg agcgcggagc gcagccactc gccgctgccc agggagcgcc 60
caagatgtgg ggggaccggg gcggcagcgg ccgtagcagc gccagggacg ggggcacgca 120
gcagcctccg ctcgcccgcc tgtcctgacc tgcctcgctt gcccccaaag aatgtcagcc 180
aagtccaagg ggaacccctc ctcgtcctgt ccagccgagg gaccgccggc agcctccaaa 240
accaaggtga aggaacagat caagatcatc gtggaggatt tggaattagt cctgggcgac 300
ctgaaggacg tggccaagga acttaaggag atgaagtccc actctgttgc ccaggctaga 360
gtgcaatggc acaatctggg ctcactgcaa cctctgcctc ccaggttcaa gctattctcc 420
tgcctcagcc tgcctcagtg cgccactacg cctgggtggt tgaccagatt gacaccctga 480
cctctgacct acagctggag gatgagatga ctgacagctc caaaacggac acgctgaata 540
gtagctcaag tggcacaaca gcctccagcc tagagaagat caaagtgcag gctaatgcac 600
cgcttattaa acccccagca cacccatctg ctatcctcac ggtcctgaga aagccaaacc 660
ctccaccacc tcctccaagg ttgacacctg tgaagtgtga agaccccaaa agggtggttc 720
caactgccaa tcctgtaaaa accaatggca cccttctacg aaatggaggc ttaccaggtg 780
gacctaacaa aattccaaat ggagatatct gctgcatacc caacagtaac ttggacaagg 840
ctccagtcca gcttctgatg catagacctg aaaaagacag atgtccccag gcagggcctc 900
gagaacgagt tcggtttaat gaaaaagtac agtaccatgg ctattgtcct gactgtgata 960
cccggtataa cataaaaaac agggaggtcc acttacacag tgaacctgtc cacccaccgg 1020
gaaagattcc tcaccaaggc cctcccctcc ctcctacacc ccatctccct cctttcccac 1080
tagaaaatgg gggaatggga ataagccaca gtaacagctt ccccctatc agacctgcaa 1140
ctgtgcctcc tcccactgca ccaaaaccac agaagacgat cttgaggaag tcaaccacta 1200
caaccgtgtg atgtatgcca ttaaaaaaat tgtttttta attttctata ttataaacat 1260
aaaataagta atgagcactt tctactcaag caataaaaag cccaaatata ttaatcctgc 1320
attcagcaaa gtggcataaa aatcacctgg taagtatgca gcacattgct tatatcctgg 1380
gtatgcatta ttttaaatgt tgtatcatta aaaacctcag aatgatgaaa aatatgaatg 1440
atgcattgtt tttgcaattg acctatgaca aactgtgaac ctgcagattt cacctatttt 1500
gatttactat aagagctggg atttgattca ttttatttat gcctaagtca tctatgcatt 1560
aacatgtcat attcttaact ttgatctaat gctttttact aggaaatttt aatactgaag 1620
gactatttta ttatttttt ctaaagatgt ttgtcactag tttttcatta ttaaatgctg 1680
aggccaatac caagaagttt attttctata ttatacaatt atgaattaca tgctcagcta 1740
tatatgtaat aaaatacttt ggtctgtgga aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800
aaaaaaaa 1808

<210> 12
<211> 281
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative protein derived from ORF of cDNA of SEQ ID NO: 11.

<400> 12

```
Met Ala Gln Ser Gly Leu Thr Ala Thr Ser Ala Ser Gln Val Gln Ala
  1               5                  10                  15

Ile Leu Leu Pro Gln Pro Ala Ser Val Arg His Tyr Ala Trp Val Val
             20                  25                  30

Asp Gln Ile Asp Thr Leu Thr Ser Asp Leu Gln Leu Glu Asp Glu Met
         35                  40                  45

Thr Asp Ser Ser Lys Thr Asp Thr Leu Asn Ser Ser Ser Ser Gly Thr
     50                  55                  60

Thr Ala Ser Ser Leu Glu Lys Ile Lys Val Gln Ala Asn Ala Pro Leu
 65                  70                  75                  80

Ile Lys Pro Pro Ala His Pro Ser Ala Ile Leu Thr Val Leu Arg Lys
                 85                  90                  95

Pro Asn Pro Pro Pro Pro Pro Pro Arg Leu Thr Pro Val Lys Cys Glu
            100                 105                 110

Asp Pro Lys Arg Val Val Pro Thr Ala Asn Pro Val Lys Thr Asn Gly
        115                 120                 125

Thr Leu Leu Arg Asn Gly Gly Leu Pro Gly Gly Pro Asn Lys Ile Pro
    130                 135                 140

Asn Gly Asp Ile Cys Cys Ile Pro Asn Ser Asn Leu Asp Lys Ala Pro
145                 150                 155                 160

Val Gln Leu Leu Met His Arg Pro Glu Lys Asp Arg Cys Pro Gln Ala
                165                 170                 175

Gly Pro Arg Glu Arg Val Arg Phe Asn Glu Lys Val Gln Tyr His Gly
            180                 185                 190

Tyr Cys Pro Asp Cys Asp Thr Arg Tyr Asn Ile Lys Asn Arg Glu Val
        195                 200                 205

His Leu His Ser Glu Pro Val His Pro Pro Gly Lys Ile Pro His Gln
    210                 215                 220

Gly Pro Pro Leu Pro Pro Thr Pro His Leu Pro Pro Phe Pro Leu Glu
225                 230                 235                 240

Asn Gly Gly Met Gly Ile Ser His Ser Asn Ser Phe Pro Pro Ile Arg
                245                 250                 255

Pro Ala Thr Val Pro Pro Pro Thr Ala Pro Lys Pro Gln Lys Thr Ile
            260                 265                 270

Leu Arg Lys Ser Thr Thr Thr Thr Val
        275                 280
```

<210> 13
<211> 1498
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived from human mesenchymal stem cells after treatment with growth supplement

<400> 13
ggcacgagcg cgactcgggc tccggacccg ggcactgctg gcggctggag cggagcgcac 60
cgcggcggtg gtgcccagag cggagcgcag ctccctgccc cgcccctccc cctcggcctc 120
gcggcgacgg cggcggtggc ggcttggacg actcggagag ccgagtgaag acatttccac 180
ctggacacct gaccatgtgc ctgccctgag cagcgaggcc caccaggcat ctctgttgtg 240
ggcagcaggg ccaggtcctg gtctgtggac cctcggcagt tggcaggctc cctctgcagt 300
ggggtctggg cctcggcccc accatgtcga gcctcggcgg tggctcccag gatgccggcg 360
gcagtagcag cagcagcacc aatggcagcg gtggcagtgg cagcagtggc ccaaaggcag 420
gagcagcaga caagagtgca gtggtggctg ccgccgcacc agcctcagtg gcagatgaca 480
caccaccccc cgagcgtcgg aacaagagcg gtatcatcag tgagcccctc aacaagagcc 540
tgcgccgctc ccgcccgctc tcccactact cttcttttgg cagcagtggt ggtagtggcg 600
gtggcagcat gatgggcgga gagtctgctg acaaggccac tgcggctgca gccgctgcct 660
ccctgttggc caatgggcat gacctggcgg cggccatggc ggtggacaaa agcaacccta 720
cctcaaagca caaaagtggt gctgtggcca gcctgctgag caaggcagag cgggccacgg 780
agctggcagc cgagggacag ctgacgctgc agcagtttgc gcagtccaca gagatgctga 840
agcgcgtggt gcaggagcat ctcccgctga tgagcgaggc gggtgctggc ctgcctgaca 900
tggaggctgt ggcaggtgcc gaagccctca atggccagtc cgacttcccc tacctgggcg 960
ctttccccat caacccaggc ctcttcatta tgaccccggc aggtgtgttc ctggccgaga 1020
gcgcgctgca catggcgggc ctggctgagt accccatgca gggagagctg gcctctgcca 1080
tcagctccgg caagaagaag cggaaacgct gcggcatgtg cgcgccctgc cggcggcgca 1140
tcaactgcga gcagtgcagc agttgtagga atcgaaagac tggccatcag atttgcaaat 1200
tcagaaaatg tgaggaactc aaaaagaagc cttccgctgc tctggagaag gtgatgcttc 1260
cgacgggagc cgccttccgg tggtttcagt gacggcggcg gaacccaaag ctgccctctc 1320
cgtgcaatgt cactgctcgt gtggtctcca gcaagggatt cgggcgaaga caaacggatg 1380
cacccgtctt tagaaccaaa aatattctct cacagatttc attcctgttt ttatatatat 1440
atttttgtt gtcgttttaa catctccacg tccctagcat aaaaaaaaaa aaaaaaaa 1498

<210> 14
<211> 361
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative protein derived from ORF of cDNA of SEQ ID NO: 13.

<400> 14
Met Ser Ser Leu Gly Gly Gly Ser Gln Asp Ala Gly Gly Ser Ser Ser
1 5 10 15

Ser Ser Thr Asn Gly Ser Gly Gly Ser Gly Ser Ser Gly Pro Lys Ala
20 25 30

Gly Ala Ala Asp Lys Ser Ala Val Val Ala Ala Ala Ala Pro Ala Ser
35 40 45

Val Ala Asp Asp Thr Pro Pro Pro Glu Arg Arg Asn Lys Ser Gly Ile
50 55 60

Ile Ser Glu Pro Leu Asn Lys Ser Leu Arg Arg Ser Arg Pro Leu Ser
65 70 75 80

His Tyr Ser Ser Phe Gly Ser Ser Gly Gly Ser Gly Gly Gly Ser Met
85 90 95

```
Met Gly Gly Glu Ser Ala Asp Lys Ala Thr Ala Ala Ala Ala Ala Ala
            100                 105                 110

Ser Leu Leu Ala Asn Gly His Asp Leu Ala Ala Ala Met Ala Val Asp
        115                 120                 125

Lys Ser Asn Pro Thr Ser Lys His Lys Ser Gly Ala Val Ala Ser Leu
    130                 135                 140

Leu Ser Lys Ala Glu Arg Ala Thr Glu Leu Ala Ala Glu Gly Gln Leu
145                 150                 155                 160

Thr Leu Gln Gln Phe Ala Gln Ser Thr Glu Met Leu Lys Arg Val Val
                165                 170                 175

Gln Glu His Leu Pro Leu Met Ser Glu Ala Gly Ala Gly Leu Pro Asp
            180                 185                 190

Met Glu Ala Val Ala Gly Ala Glu Ala Leu Asn Gly Gln Ser Asp Phe
        195                 200                 205

Pro Tyr Leu Gly Ala Phe Pro Ile Asn Pro Gly Leu Phe Ile Met Thr
    210                 215                 220

Pro Ala Gly Val Phe Leu Ala Glu Ser Ala Leu His Met Ala Gly Leu
225                 230                 235                 240

Ala Glu Tyr Pro Met Gln Gly Glu Leu Ala Ser Ala Ile Ser Ser Gly
                245                 250                 255

Lys Lys Lys Arg Lys Arg Cys Gly Met Cys Ala Pro Cys Arg Arg Arg
            260                 265                 270

Ile Asn Cys Glu Gln Cys Ser Ser Cys Arg Asn Arg Lys Thr Gly His
        275                 280                 285

Gln Ile Cys Lys Phe Arg Lys Cys Glu Glu Leu Lys Lys Lys Pro Ser
    290                 295                 300

Ala Ala Leu Glu Lys Val Met Leu Pro Thr Gly Ala Ala Phe Arg Trp
305                 310                 315                 320

Phe Gln Lys Thr Lys Ile Leu Lys Glu Met Ser Gly Gln Arg Gln Gly
                325                 330                 335

Pro Gly His Val Ala Cys Pro Ala Ser Met Ser Leu Gly Gly Gly Val
            340                 345                 350

Gly Trp Gly Trp Ala Ala Ala Ser Gln
        355                 360
```

<210> 15
<211> 2329
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived
from human mesenchymal stem cells after treatment

with growth supplement

<400> 15

ggcacgagat gacattactc actatcagaa ttgagaaaat tggtttgaaa gatgctgggc 60
agtgcatcga tccctatatt acagttagtg taaaggatct gaatggcata gacttaactc 120
ctgtgcaaga tactcctgtg gcttcaagaa aagaagatac atatgttcat tttaatgtgg 180
acattgagct ccagaagcat gttgaaaaat taaccaaagg tgcagctatc ttctttgaat 240
tcaaacacta caagcctaaa aaaaggttta ccagcaccaa gtgttttgct ttcatggaga 300
tggatgaaat taaacctggg ccaattgtaa tagaactata caagaaaccc actgacttta 360
aaagaaagaa attgcaatta ttgaccaaga aaccacttta tcttcatcta catcaaactt 420
tgcacaagga atgatcctga catgatgaac ctggaacttc tgtgaatttt accactcagt 480
agaaaccatc atagctctgt gtagcatatt caccttcaa caggcaggaa gcaagccgta 540
cccagaccag taggccggac ggagtccaat gcaaagctgt accacagaat tcagagtcca 600
gcacatcaca ctgacgtata ggactccttg ggatacaggt ttattgtaga ttttgaaaca 660
tgtttttact tttctattaa ttgtgcaatt aatagtctat tttctaattt accactactc 720
ctaccctgct tcctggaaca atactgttgt gggtaggatg tgctcatctt cagacttaat 780
acagcaataa gaatgtgcta gagtttacac atctgttcac ttttgctcca atatgctctt 840
ttgacttaac gtcaagcttt gggttgatgt gggtagggta gtgtcaaact gctttgagag 900
gaatgggacc agttctgctg cctaagaagg tctgtctgga tgtttatagg cagcacctct 960
gaagtggcct aaattcaccc tgatctgata gttttcctgc ttagaaagtg tgccttggcc 1020
agatcagtat cccacatggg agtgttccct aggttgtagc tgtgattgtt tccagatgac 1080
cagattgttt ttctgaaaat gagcatattt ttagtcatgt cgattagctg ttcttctaca 1140
tcacattgtt actctttctg atgatgattc tagggttaac attggaacca tctcaaaata 1200
attacaaagt tttagatggg tttacaatgt cttctaaaca atgtaatcta aaaataattg 1260
agtcagatgc taacgagata ctgcaggcat aactgctgtt tttctgacaa ctgattgtga 1320
aaccttaaaa cctgcatacc tcttcttaca gtgaggagta tgcaaaatct ggaaagatat 1380
tctatttttt ttatataggt agataggatc gccatttatt tcctatttag atatactgac 1440
attcatccat atgaaaatat gcaggtcatt agcttactat aatttacttt tgacttaatg 1500
gggcataaat aaaactttca tagtacacga ggtggatatt tgatacacag aacatttgcg 1560
gtgggctttc tgtgggttag atgtaaagcc cacatatttt aatattcact attttaaatg 1620
agcaatgcat gaggggaatg cagtgtcagt acctggccta tttttaaact agtgtaatca 1680
ccctagtcat accattcagt atgtttgctt tttaaaataa gtaaccacaa ttaagttgtt 1740
gtagcccttg cacttcaaga gatctagtct ttactttcag ttgtctgtta ggtccattct 1800
gtttactaga cggatgttaa taaaaactat gcgagcctga atgaattctc agccaaattt 1860
agtcttgtct ctcatcttga ttggattaat tccaaattct aaaatgattc agtccacaat 1920
agctctaggg gatgaagaat ttgccttact ttgcccagtt cctaagactg tgagttgtca 1980
aatccctaga ctgtaagctc ttcaaggagc aagaggcgca ttttctccgt gtcatgtaat 2040
ttttctaagg tgcttggcag cactctgacc ctgtggagta ctcagacctt ttgtttgatg 2100
ttgctgacaa gacctgaaaa aaaatccctt aaaaaaaaaa acccattaaa gtgtagcaaa 2160
accgatgtat tatttttttct cttttctcat aaaccataag atatgtatat caaaatgata 2220
atttatatga agcatctgct atccttcaca taagtatagt ttatatgaac tacaaatacc 2280
tttcatcaac attaaagaag cttaaattga aaaaaaaaaa aaaaaaaaa 2329

<210> 16
<211> 141
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative
protein derived from ORF of cDNA of SEQ ID NO: 15.

<400> 16

Met Thr Leu Leu Thr Ile Arg Ile Glu Lys Ile Gly Leu Lys Asp Ala
1 5 10 15

Gly Gln Cys Ile Asp Pro Tyr Ile Thr Val Ser Val Lys Asp Leu Asn
20 25 30

Gly Ile Asp Leu Thr Pro Val Gln Asp Thr Pro Val Ala Ser Arg Lys

```
         35                  40                  45

Glu Asp Thr Tyr Val His Phe Asn Val Asp Ile Glu Leu Gln Lys His
      50                  55                  60

Val Glu Lys Leu Thr Lys Gly Ala Ala Ile Phe Phe Glu Phe Lys His
 65                  70                  75                  80

Tyr Lys Pro Lys Lys Arg Phe Thr Ser Thr Lys Cys Phe Ala Phe Met
                 85                  90                  95

Glu Met Asp Glu Ile Lys Pro Gly Pro Ile Val Ile Glu Leu Tyr Lys
            100                 105                 110

Lys Pro Thr Asp Phe Lys Arg Lys Lys Leu Gln Leu Leu Thr Lys Lys
        115                 120                 125

Pro Leu Tyr Leu His Leu His Gln Thr Leu His Lys Glu
    130                 135                 140
```

<210> 17
<211> 1737
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived
from human mesenchymal stem cells after treatment
with growth supplement

<400> 17

```
ggcacgagtc gacatggcgg ttaccctgag tctcttgctg ggcgggcgcg tttgcgccgc 60
cgtcactcgc tgtgggttcg cgacccgggg ggtggcgggc ccaggcccta ttggccggga 120
gccggacccc gattccgact gggagccgga ggaacgggag ctgcaggagg tggagaggta 180
ccggcttctc cccgggccct cagcttgaag cagggcctcg tgccccggcg ctccaggccg 240
cgccccttg gcgccgggtg tcctgccgct gcttgcgcag cggcccttgt ttcttcttac 300
ccgttgttag gggcgcagcg tcaggtgttc agctccccgg gaccactggt cccttcatta 360
gtgacgtatt tcatcatcag tttagagagt tcggcatgct tacaggcagt tattgttcta 420
ggtgttagct ttctgggtgt acggagcagc tctaagccgg caacatggcc cggttgccct 480
tgcgatcaaa gagaagaggg ctgggcgctc catgatttag cctgaggctc ttcaaacatc 540
cattctgctt ccacgcatgg cttctgccat tggttctctt cccccagcac cctgaaacga 600
cagaaacaag caatccgatt ccagaaaatt cggaggcaaa tggaggcgcc tggtgccccg 660
cccaggaccc tgacgtggga agccatggag cagatacggt atttacatga ggaatttcca 720
gagtcctggt cagttcccag gttggctgaa ggctttgatg tcagcactga tgtgatccga 780
agagttttaa aaagcaagtt tttacccaca ttggagcaga agctgaagca ggatcaaaaa 840
gtccttaaga aagctgggct tgcccactcg ctgcagcacc tccggggctc tggaaatacc 900
tcaaagctgc tccctgcagg ccactctgta tcaggctctt tgcttatgcc agggcatgaa 960
gcctcatcta aagacccaaa tcacagcaca gctttgaaag tgatagagtc agacactcac 1020
aggacaaata caccaaggag aaggaaggga agaaataaag aaatccagga cctggaggag 1080
agctttgtgc ctgttgctgc acccctaggt catccaagag agctgcagaa gtactccagt 1140
gattctgaga gccccagagg aactggcagt ggtgcgttgc caagtggtca gaagctggag 1200
gagttgaagg cagaggagcc agataacttc agcagcaaag tagtgcagag gggccgagag 1260
ttctttgaca gcaacgggaa cttcctgtac agaatttgag tcggggcttg gcttatggag 1320
atgcctcgtg aaacacagct gggcaagtat taatgtatat ggaacagcct ggatttctgc 1380
atatggataa gccaccttgg aataggaaga ggtgttgagc ctggactgtg ggaggaaaga 1440
gctgcgtgga tagattcaaa cttcctgtgg tagtgctccc agtctgacct ctgtagacct 1500
tcagtactca ctcttcttgc ttaggctctc tgtgtgttga aagccatccc gtgttgcatg 1560
tgttgttaca attttctgtg atacttgcaa tttatgtttg agaagaagtg aaaagtttgc 1620
cttctgacct catttccttc ttgatcagtg aacactaaca ttttggggac aacttagtca 1680
attggttttc cttacaacaa aataaagtaa aatgtagcaa aaaaaaaaaa aaaaaaa    1737
```

<210> 18
<211> 219
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative protein derived from ORF of cDNA of SEQ ID NO: 17.

<400> 18

```
Met Glu Ala Pro Gly Ala Pro Pro Arg Thr Leu Thr Trp Glu Ala Met
  1               5                  10                  15

Glu Gln Ile Arg Tyr Leu His Glu Glu Phe Pro Glu Ser Trp Ser Val
             20                  25                  30

Pro Arg Leu Ala Glu Gly Phe Asp Val Ser Thr Asp Val Ile Arg Arg
         35                  40                  45

Val Leu Lys Ser Lys Phe Leu Pro Thr Leu Glu Gln Lys Leu Lys Gln
     50                  55                  60

Asp Gln Lys Val Leu Lys Lys Ala Gly Leu Ala His Ser Leu Gln His
 65                  70                  75                  80

Leu Arg Gly Ser Gly Asn Thr Ser Lys Leu Leu Pro Ala Gly His Ser
                 85                  90                  95

Val Ser Gly Ser Leu Leu Met Pro Gly His Glu Ala Ser Ser Lys Asp
            100                 105                 110

Pro Asn His Ser Thr Ala Leu Lys Val Ile Glu Ser Asp Thr His Arg
        115                 120                 125

Thr Asn Thr Pro Arg Arg Arg Lys Gly Arg Asn Lys Glu Ile Gln Asp
    130                 135                 140

Leu Glu Glu Ser Phe Val Pro Val Ala Ala Pro Leu Gly His Pro Arg
145                 150                 155                 160

Glu Leu Gln Lys Tyr Ser Ser Asp Ser Glu Ser Pro Arg Gly Thr Gly
                165                 170                 175

Ser Gly Ala Leu Pro Ser Gly Gln Lys Leu Glu Glu Leu Lys Ala Glu
            180                 185                 190

Glu Pro Asp Asn Phe Ser Ser Lys Val Val Gln Arg Gly Arg Glu Phe
        195                 200                 205

Phe Asp Ser Asn Gly Asn Phe Leu Tyr Arg Ile
    210                 215
```

<210> 19
<211> 369
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:cDNA derived
from human mesenchymal stem cells after treatment
with growth supplement

<400> 19
cattaggtgt caccaaagtg ttaacaattc cttgggaatt tataaacaaa aaatatctag 60
actaaaaata gaatataagg cccctctgga accactgcac acctttccct ctgttctcag 120
attaaactgc tgcataaaat gaggttattt agcctgtttt taaggaaggc ttctattgca 180
catgcatctc tttggttggt aataacatgt acttttgtta aaatatttcc ataatgaagc 240
cctgttgtgt ggttagctgg gtgtggactt tcctcccttt cttgggggcc ctcctactca 300
cagtcaagtt gccctttaga actaaagatc tggtaggatt gggtctttat tcaatagccc 360
ctcgtgccg 369

<210> 20
<211> 48
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative
protein derived from ORF of cDNA of SEQ ID NO: 19.

<400> 20
Glu Ser Ile Arg Cys His Gln Ser Val Asn Asn Ser Leu Gly Ile Tyr
1 5 10 15

Lys Gln Lys Ile Ser Arg Leu Lys Ile Glu Tyr Lys Ala Pro Leu Glu
20 25 30

Pro Leu His Thr Phe Pro Ser Val Leu Arg Leu Asn Cys Cys Ile Lys
35 40 45

<210> 21
<211> 2133
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived
from human mesenchymal stem cells after treatment
with growth supplement

<400> 21
cggcacgaga aatcatttaa tttaaggatt gcagcttaag acttggcata cggatacttg 60
gatgacactg atttcttctg tagctgggca aatcagtttc ctcaccaaaa gtgaatatgg 120
taaaattaca tgtttatatc atatgattgt gagactcaat aatagtttga gagctcagta 180
ttatcttttgt aggaaagaca taatttccaa atttgattag aatttgaaaa taaatcacat 240
tttgatcatt gctagatact gtgtctaaat gtggaaaaac ctttgagaag gatttttgta 300
ttttttttttg actatttcaa gtcaatagag aaaagaattc tgttgaaatc agtgaagtgg 360
tgataagtag tgggtggaag attacagctc agacttaatg ggatagattt ggaatattat 420
agatgggtga atttattaca aagcaaatat taccccaaat tgtcctctaa gaaaatcctc 480
ccccaatctg ttaatgaaaa aaaaattgtg gtacaatcct tataaaattt accattttag 540
ccattttaaa gtcttaatgc cagttcagtg gcattaagta cattcatttt gctttgttac 600
tatcactgcc atccatctcc agaaagcttt tcatcttgaa aatctgttac cgtacttatt 660
taacaataag tccttatatc cctctctcct cagcccctgg caaccaccat cctactttct 720
atctctgaat tcgactcttc tattaatagg tacctcatat aagtggaaca tagagtattt 780

gccctttttt gtgactggct tatttcactt agcataatgt gcccaagagt catatatgtt 840
agcatgcgtc aggattttat ttcttttaaa ggctaaataa tgttccattg tatgtatata 900
ccatattttg cttattcttt caccttccat ggacatttag gttgtttcta ccttgtaggt 960
actgtgaatc atgcagccat aaacattggt gttcaaatat atgttccagc ttctgctttc 1020
aattcttttg catatatacc cagaagttaa attgctggat cacatagtca ttctattttt 1080
aatttttttg aggaactgtc atattgtttt ctatagtggc tgcaccacca aaattttgaa 1140
tggagagaat gtcagcatcc tgtttaataa attattttaa aaatcagctc tgagcaaatg 1200
aatttagaat ttggaatatt tgaccaaaaa tcagttttac acgtacatag tatgttactg 1260
ataaaggaat gactagagta gtacacctac agttcttgct ggccagagcc aggagggggt 1320
gtaaaggaat ttcctgggtt taaaaaagca ggcagcattg atgaatcatg aagttccccc 1380
tacctccttt aatggttttg taacaaagag atgcattcag gtgaaagttt caagtctttt 1440
tgtaactata gaataactgt tctttttgta attgtagaat tctttaactt tttttaggct 1500
gaagtagaga acatgaatca tttggtctga aacttttgag accctatttg gtttgcttac 1560
atatgtgaaa atagttatcc aaataattgc gattgctatc aaataatttc aataatatca 1620
gttacttctt ttggtctaga gtcttctcct ccttaacctc tttctttttg gcgagacagg 1680
gtctcactct gttacccagg ctggagtgca atggagtgat cactgcccct ttgaacttct 1740
gggctcaagt gatcctctca cctcagcttc ctgagtaggt gggactacag atgtgtgcca 1800
ccaagcctgg ctaattaaaa acaaatgttt aagccaggca tggtggctca cacctgtaat 1860
ttcagcacct tgggaggccg aggcaggcgg atcacttgag gtcaggagtt caagaccagc 1920
ctggccaaga tggtggaacc ctgcctctac taaaaacaca cacacaaatt agccgggcat 1980
ggtggctggt gtctgtaatc ccagctgtgg gaggttgagg caggagaatc acttgaacct 2040
gggaggtgga ggtttcagtg agccaagatc atttcacctc actccagcct gggtgacagc 2100
gtaagactct ctcaaaaaaa aaaaaaaaaa aaa 2133

<210> 22
<211> 56
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative
protein derived from ORF of cDNA of SEQ ID NO: 21.

<400> 22
Met Phe His Cys Met Tyr Ile Pro Tyr Phe Ala Tyr Ser Phe Thr Phe
1 5 10 15

His Gly His Leu Gly Cys Phe Tyr Leu Val Gly Thr Val Asn His Ala
20 25 30

Ala Ile Asn Ile Gly Val Gln Ile Tyr Val Pro Ala Ser Ala Phe Asn
35 40 45

Ser Phe Ala Tyr Ile Pro Arg Ser
50 55

<210> 23
<211> 1200
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived
from human mesenchymal stem cells after treatment
with growth supplement

<400> 23
ggcacgaggt agcagcagca tggccgcgat ctatgggggt gtagaggggg gaggcacacg 60
atccgaggtc cttttagtct cagaggatgg gaagatcctg gcagaagcag atggactgag 120

cacaaaccac tggctgatcg ggacagacaa gtgtgtggag aggatcaatg agatggtgaa 180
cagggccaaa cggaaagcag gggtggatcc tctggtaccg ctgcgaagct tgggcctatc 240
tctgagcggt ggggaccagg aggacgcggg gaggatcctg atcgaggagc tgagggaccg 300
atttccctac ctgagtgaaa gctacttaat caccaccgat gccgccggct ccatcgccac 360
agctacaccg gatggtggag ttgtgctcat atctggaaca ggctccaact gcaggctcat 420
caaccctgat ggctccgaga gtggctgcgg cggctggggc catatgatgg gtgatgaggg 480
ttcagcctac tggatcgcac accaagcagt gaaaatagtg tttgactcca ttgacaacct 540
agaggcggct cctcatgata tcggctacgt caaacaggcc atgttccact atttccaggt 600
gccagatcgg ctagggatac tcactcacct gtatagggac tttgataaat gcaggtttgc 660
tgggttttgc cggaaaattg cagaaggtgc tcagcaggga gacccccttt cccgctatat 720
cttcaggaag gctggggaga tgctgggcag acacatcgta gcagtgttgc ccgagattga 780
cccggtcttg ttccagggca agattggact ccccatcctg tgcgtgggct ctgtatggaa 840
gagctgggag ctgctgaagg aaggttttct tttggcgctg acccagggca gagagatcca 900
ggctcagaac ttcttctcca gcttcaccct gatgaagctg aggcactcct ccgctctggg 960
tggggccagc ctaggggcca ggcacatcgg gcacctcctc cccatggact atagcgccaa 1020
tgccattgcc ttctattcct acacctttc ctaggggct ggtcccggct ccacccctc 1080
caagctcagt ggacactggg tctgaaagga aggagtcttt tgcttccttt ctccttttta 1140
caaaaacaaa catagaagaa aataaatgca ctttatccaa aaaaaaaaaa aaaaaaaaaa 1200

<210> 24
<211> 344
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative
protein derived from ORF of cDNA of SEQ ID NO: 23.

<400> 24
Met Ala Ala Ile Tyr Gly Gly Val Glu Gly Gly Gly Thr Arg Ser Glu
1 5 10 15

Val Leu Leu Val Ser Glu Asp Gly Lys Ile Leu Ala Glu Ala Asp Gly
20 25 30

Leu Ser Thr Asn His Trp Leu Ile Gly Thr Asp Lys Cys Val Glu Arg
35 40 45

Ile Asn Glu Met Val Asn Arg Ala Lys Arg Lys Ala Gly Val Asp Pro
50 55 60

Leu Val Pro Leu Arg Ser Leu Gly Leu Ser Leu Ser Gly Gly Asp Gln
65 70 75 80

Glu Asp Ala Gly Arg Ile Leu Ile Glu Glu Leu Arg Asp Arg Phe Pro
85 90 95

Tyr Leu Ser Glu Ser Tyr Leu Ile Thr Thr Asp Ala Ala Gly Ser Ile
100 105 110

Ala Thr Ala Thr Pro Asp Gly Gly Val Val Leu Ile Ser Gly Thr Gly
115 120 125

Ser Asn Cys Arg Leu Ile Asn Pro Asp Gly Ser Glu Ser Gly Cys Gly
130 135 140

Gly Trp Gly His Met Met Gly Asp Glu Gly Ser Ala Tyr Trp Ile Ala
145 150 155 160

His Gln Ala Val Lys Ile Val Phe Asp Ser Ile Asp Asn Leu Glu Ala

```
                165                 170                 175

Ala Pro His Asp Ile Gly Tyr Val Lys Gln Ala Met Phe His Tyr Phe
            180                 185                 190

Gln Val Pro Asp Arg Leu Gly Ile Leu Thr His Leu Tyr Arg Asp Phe
        195                 200                 205

Asp Lys Cys Arg Phe Ala Gly Phe Cys Arg Lys Ile Ala Glu Gly Ala
    210                 215                 220

Gln Gln Gly Asp Pro Leu Ser Arg Tyr Ile Phe Arg Lys Ala Gly Glu
225                 230                 235                 240

Met Leu Gly Arg His Ile Val Ala Val Leu Pro Glu Ile Asp Pro Val
                245                 250                 255

Leu Phe Gln Gly Lys Ile Gly Leu Pro Ile Leu Cys Val Gly Ser Val
            260                 265                 270

Trp Lys Ser Trp Glu Leu Leu Lys Glu Gly Phe Leu Leu Ala Leu Thr
        275                 280                 285

Gln Gly Arg Glu Ile Gln Ala Gln Asn Phe Phe Ser Ser Phe Thr Leu
    290                 295                 300

Met Lys Leu Arg His Ser Ser Ala Leu Gly Gly Ala Ser Leu Gly Ala
305                 310                 315                 320

Arg His Ile Gly His Leu Leu Pro Met Asp Tyr Ser Ala Asn Ala Ile
                325                 330                 335

Ala Phe Tyr Ser Tyr Thr Phe Ser
            340
```

<210> 25
<211> 2216
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:cDNA derived
from human mesenchymal stem cells after treatment
with growth supplement

<400> 25

```
ggcacgaggg gcaccgtaac cagcgccgcg gacaccggca ccggcgccac ggactccgca 60
ggaccccgcg cccgccgccg ccgctatgct ggggctgctg gtggcgttgc tggccctggg 120
gctcgctgtc tttgcgctgc tggacgtctg gtacctggtg cgccttccgt gcgccgtgct 180
gcgcgcgcgc ctgctgcagc cgcgcgtccg tgacctgcta gctgagcagc gcttcccggg 240
ccgcgtgctg ccctcggact tggacctgct gttgcacatg aacaacgcgc gctacctgcg 300
cgaggccgac tttgcgcgcg tcgcgcacct gacccgctgc ggggtgctcg gggcgctgag 360
ggagttgcgg gcgcacacgg tgctggcggc ctcgtgcgcg cgccaccgcc gctcgctgcg 420
cctgctggag cccttcgagg tgcgcacccg cctgctgggc tgggacgacc gcgcgttcta 480
cctggaggcg cgctttgtca gcctgcggga cggtttcgtg tgcgcgctgc tgcgcttccg 540
gcagcacctg ctgggcacct cacccgagcg cgtcgtgcag cacctgtgcc agcgcagggt 600
ggagcccccc gagctgcccg ctgatctgca gcactggatc tcctacaacg aggccagcag 660
ccagctgctc cgcatggaga gtgggctcag tgatgtcacc aaggaccagt gaccgccacc 720
ttcacaccgt ctgccctggc caccatcctg ggcctggggg ctgcccacag atgggcagtc 780
tcagccatac tctgttccag ctggagtagc ctcctgacca gcctggccca ccctgctcca 840
```

cccactgggc ccccccagtt attgataccc ctctgtgctg ggctccacgc taggcagaag 900
gaggagtggc attggcatcc tgacccagct ctgccctcaa ggtggggatg gatgggcaaa 960
ggagagtcct gcctggccct acgatgaggc cactcatgtg ggcctaggta ggggaggatg 1020
gtgcctggag cagagggacc cacaagtgcc tcccgagcct agatcctggc tcggaccact 1080
gcaagggccg aggcagggcc agaccagagc atcctgggta caggcctggg ctctccaggg 1140
cctgggcctg attcaggtgc agtgggcact cctgaagggt cagagcggca tctgccaggc 1200
agcccctctg gcttccgctg aggtggttgc aggcctgggg cagagcctgg gtggtcagag 1260
gccggggcta gaggcagatg gaagggaggc atttgctgac agaggacggg gcacccgggc 1320
tccccactgc agtcggcctt gcctcctcct cctcctctac ctccagtcag gctggacggg 1380
agggtagcct tgtggctgag aggggtcaga ctaggtggca caggggctcc tggaaagaca 1440
gcaggcttcc tgctgggcgt tcccttgttg gagggaatag agtgggggtg ggactctgca 1500
ggggtgtcct tgtccactcg cacccctcgc cgcccaccag ggccatgctc tgtgacttgg 1560
gctgatcccc acccttttctg ggcctacagc accacaggcc gctgtacccc cttagagctg 1620
cccctctctg gcctggccgg caggcgtctt cttaactcct ctgtcctcta tattcagcat 1680
gttccttgtc agctgctggg ccggccctgc cttgcgctag cagagcctct cctggcagct 1740
tctcaggtct ccctaatgga gacaccaggc tactaggaca ctggctgggg ccacccccctc 1800
ctgcctaatg cctcacctta cagctgggga aactgaggcc tggaatggcc cagagtcacc 1860
aaggcaaagt tggggctggt cccagcctga ggctccagct gatgccctca gctcccagag 1920
agggggtgcc ccatctagct gggtgcaggg gtcactgctt gtcagctcag ggccctgtgc 1980
ccgcttgcct gttcccctac atctgtgcct gcacatccag aactgcctcc ttgccgctgc 2040
ctccaggaag cccaccttga gccagagtca agggctgcag cactgcccga tagaacacgc 2100
ccgccctcac tgctgttctt gccttacagc caccatggga aagctgcaac ctttctgttt 2160
tatttaaaga aagcccaaca ttaaagggtt ttcattgcaa aaaaaaaaaa aaaaaa 2216

<210> 26
<211> 208
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Putative
protein derived from ORF of cDNA of SEQ ID NO: 25.

<400> 26
Met Leu Gly Leu Leu Val Ala Leu Leu Ala Leu Gly Leu Ala Val Phe
1 5 10 15

Ala Leu Leu Asp Val Trp Tyr Leu Val Arg Leu Pro Cys Ala Val Leu
20 25 30

Arg Ala Arg Leu Leu Gln Pro Arg Val Arg Asp Leu Leu Ala Glu Gln
35 40 45

Arg Phe Pro Gly Arg Val Leu Pro Ser Asp Leu Asp Leu Leu Leu His
50 55 60

Met Asn Asn Ala Arg Tyr Leu Arg Glu Ala Asp Phe Ala Arg Val Ala
65 70 75 80

His Leu Thr Arg Cys Gly Val Leu Gly Ala Leu Arg Glu Leu Arg Ala
85 90 95

His Thr Val Leu Ala Ala Ser Cys Ala Arg His Arg Arg Ser Leu Arg
100 105 110

Leu Leu Glu Pro Phe Glu Val Arg Thr Arg Leu Leu Gly Trp Asp Asp
115 120 125

Arg Ala Phe Tyr Leu Glu Ala Arg Phe Val Ser Leu Arg Asp Gly Phe
130 135 140

Val Cys Ala Leu Leu Arg Phe Arg Gln His Leu Leu Gly Thr Ser Pro
145 150 155 160

Glu Arg Val Val Gln His Leu Cys Gln Arg Arg Val Glu Pro Pro Glu
165 170 175

Leu Pro Ala Asp Leu Gln His Trp Ile Ser Tyr Asn Glu Ala Ser Ser
180 185 190

Gln Leu Leu Arg Met Glu Ser Gly Leu Ser Asp Val Thr Lys Asp Gln
195 200 205

<210> 27
<211> 2453
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Consensus sequence of the coding strand of cDNAs from human mesenchymal stem cell genes.

<400> 27
gaattcggca cgaggtcgcg gcggcgaagg aggaggagga agagggcgag gcgacaagag 60
aagaaggagg caggcgcggc ggcagcggcg gcgccccgag ccggcggagg cgaggggggg 120
gaagatggcg gacgtgctta gcgtcctgcg acagtacaac atccagaaga aggagattgt 180
ggtgaaggga gacgaagtga tcttcgggga gttctcctgg cccaagaatg tgaagaccaa 240
ctatgttgtt tgggggactg gaaaggaagg ccaacccaga gagtactaca cattggattc 300
cattttattt ctacttaata acgtgcacct ttctcatcct gtttatgtcc gacgtgcagc 360
tactgaaaat attcctgtgg ttagaagacc tgatcgaaaa gatctacttg gatatctcaa 420
tggtgaagcg tcaacatcgg caagtataga cagaagcgct cccttagaaa taggtcttca 480
gcgatctact caagtcaaac gagctgcaga tgaagtttta gcagaagcaa agaaaccacg 540
aattgaggat gaagagtgtg tgcgccttga taaagagaga ttggctgccc gtttggaggg 600
tcacaaagaa gggattgtac agactgaaca gattaggtct ttgtctgaag ctatgtcagt 660
ggaaaaaatt gctgcaatca aagccaaaat tatggctaag aaaagatcta ctatcaagac 720
tgatctagat gatgacataa ctgcccttaa acagaggagt tttgtggatg ctgaggtaga 780
tgtgacccga gatattgtca gcagagagag agtatggagg acacgaacaa ctatcttaca 840
aagcacagga aagaattttt ccaagaacat ttttgcaatt cttcaatctg taaaagccag 900
agaagaaggg cgtgcacctg aacagcgacc tgccccaaat gcagcacctg tggatcccac 960
tttgcgcacc aaacagccta tcccagctgc ctataacaga tacgatcagg aaagattcaa 1020
aggaaaagaa gaaacggaag gcttcaaaat tgacactatg ggaacctacc atggtatgac 1080
actgaaatct gtaacggagg gtgcatctgc ccggaagact cagactcctg cagcccagcc 1140
agtaccaaga ccagtttctc aagcaagacc tcccccaaat cagaagaaag gatctcgaac 1200
acccattatc ataattcctg cagctaccac ctctttaata accatgctta atgcaaaaga 1260
ccttctacag gacctgaaat ttgtcccatc agatgaaaag aagaaacaag gttgtcaacg 1320
agaaaatgaa actctaatac aaagaagaaa agaccagatg caaccagggg gcactgcaat 1380
tagtgttaca gtaccttata gagtagtaga ccagcccctt aaacttatgc ctcaagactg 1440
ggaccgcgtt gtagccgttt ttgtgcaggg tcctgcatgg cagttcaaag gttggccatg 1500
gcttttgcct gatggatcac cagttgatat atttgctaaa attaaagcct tccatctgaa 1560
gtatgatgaa gttcgtctgg atccaaatgt tcagaaatgg gatgtaacag tattagaact 1620
cagctatcac aaacgtcatt tggatagacc agtgttctta cggttttggg aaacattgga 1680
caggtacatg gtaaagcata aatcgcactt gagattctga attatttggc tcctccattt 1740
ctggaaattg agactcaagc tttatgaatt tatcaagaac ttaaaaatga agaaggtcac 1800
agattgatct tttataagac cttatttgat gctttgtgct tcaaggagat gatacctgtc 1860
atccatataa gcaaactttt tggcttacaa ctatttttt aatattagcc ttctagtctg 1920
taatggaaat tgtatatttt gatagaagtt ttttctccat tggttaaatt agcattactt 1980

aaaatttgtt tctttagaaa ataaatgcag gttataaatg tgtgtatatt tagagattat 2040
aaggctctct gagccatctt ctgatttttc attgctctat aattcttttt actgaaaata 2100
ctatgttatg aatggtatta aattttagtc tctggaacat ccaaaaccaa gcaaagggat 2160
gtgactattt tgaatgaatc agaatgtcaa cttgtatgta cactatatct acacttactc 2220
attatttaaa aagaataatg aaaaatctag atcaattctt caatttgatt gaactgttca 2280
gccttttcaa gatttcttta tttacaaatg attacattta aatgaatgta cattcttctc 2340
actgactttg gtgattttga aacctagaat gatgtgtttc tatctgtaat atctttccat 2400
ttgaaaaaaa tctcaaaaca cagattaaaa ccacaaaaaa aaaaaaaaaa aaa 2453

<210> 28
<211> 1593
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Coding sequence
for the protein of SEQ ID NO:29 and corresponding
to residues 125-1717 of SEQ ID NO:27

<400> 28
atggcggacg tgcttagcgt cctgcgacag tacaacatcc agaagaagga gattgtggtg 60
aagggagacg aagtgatctt cggggagttc tcctggccca agaatgtgaa gaccaactat 120
gttgtttggg ggactggaaa ggaaggccaa cccagagagt actacacatt ggattccatt 180
ttatttctac ttaataacgt gcacctttct catcctgttt atgtccgacg tgcagctact 240
gaaaatattc ctgtggttag aagacctgat cgaaaagatc tacttggata tctcaatggt 300
gaagcgtcaa catcggcaag tatagacaga agcgctccct tagaaatagg tcttcagcga 360
tctactcaag tcaaacgagc tgcagatgaa gttttagcag aagcaaagaa accacgaatt 420
gaggatgaag agtgtgtgcg ccttgataaa gagagattgg ctgcccgttt ggagggtcac 480
aaagaaggga ttgtacagac tgaacagatt aggtctttgt ctgaagctat gtcagtggaa 540
aaaattgctg caatcaaagc caaaattatg gctaagaaaa gatctactat caagactgat 600
ctagatgatg acataactgc ccttaaacag aggagttttg tggatgctga ggtagatgtg 660
acccgagata ttgtcagcag agagagagta tggaggacac gaacaactat cttacaaagc 720
acaggaaaga atttttccaa gaacattttt gcaattcttc aatctgtaaa agccagagaa 780
gaagggcgtg cacctgaaca gcgacctgcc ccaaatgcag cacctgtgga tcccactttg 840
cgcaccaaac agcctatccc agctgcctat aacagatacg atcaggaaag attcaaagga 900
aaagaagaaa cggaaggctt caaaattgac actatgggaa cctaccatgg tatgacactg 960
aaatctgtaa cggagggtgc atctgcccgg aagactcaga ctcctgcagc ccagccagta 1020
ccaagaccag tttctcaagc aagacctccc ccaaatcaga agaaaggatc tcgaacaccc 1080
attatcataa ttcctgcagc taccacctct ttaataacca tgcttaatgc aaaagacctt 1140
ctacaggacc tgaaatttgt cccatcagat gaaaagaaga aacaaggttg tcaacgagaa 1200
aatgaaactc taatacaaag aagaaaagac cagatgcaac caggggcac tgcaattagt 1260
gttacagtac cttatagagt agtagaccag ccccttaaac ttatgcctca agactgggac 1320
cgcgttgtag ccgtttttgt gcagggtcct gcatggcagt tcaaaggttg gccatggctt 1380
ttgcctgatg gatcaccagt tgatatattt gctaaaatta aagccttcca tctgaagtat 1440
gatgaagttc gtctggatcc aaatgttcag aaatgggatg taacagtatt agaactcagc 1500
tatcacaaac gtcatttgga tagaccagtg ttcttacggt tttgggaaac attggacagg 1560
tacatggtaa agcataaatc gcacttgaga ttc 1593

<210> 29
<211> 531
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:Amino acid
sequence derived from the nucleotide sequence of
SEQ ID NO:28

<400> 29

```
Met Ala Asp Val Leu Ser Val Leu Arg Gln Tyr Asn Ile Gln Lys Lys
 1               5                  10                  15

Glu Ile Val Val Lys Gly Asp Glu Val Ile Phe Gly Glu Phe Ser Trp
            20                  25                  30

Pro Lys Asn Val Lys Thr Asn Tyr Val Val Trp Gly Thr Gly Lys Glu
        35                  40                  45

Gly Gln Pro Arg Glu Tyr Tyr Thr Leu Asp Ser Ile Leu Phe Leu Leu
    50                  55                  60

Asn Asn Val His Leu Ser His Pro Val Tyr Val Arg Arg Ala Ala Thr
65                  70                  75                  80

Glu Asn Ile Pro Val Val Arg Arg Pro Asp Arg Lys Asp Leu Leu Gly
                85                  90                  95

Tyr Leu Asn Gly Glu Ala Ser Thr Ser Ala Ser Ile Asp Arg Ser Ala
            100                 105                 110

Pro Leu Glu Ile Gly Leu Gln Arg Ser Thr Gln Val Lys Arg Ala Ala
        115                 120                 125

Asp Glu Val Leu Ala Glu Ala Lys Lys Pro Arg Ile Glu Asp Glu Glu
    130                 135                 140

Cys Val Arg Leu Asp Lys Glu Arg Leu Ala Ala Arg Leu Glu Gly His
145                 150                 155                 160

Lys Glu Gly Ile Val Gln Thr Glu Gln Ile Arg Ser Leu Ser Glu Ala
                165                 170                 175

Met Ser Val Glu Lys Ile Ala Ala Ile Lys Ala Lys Ile Met Ala Lys
            180                 185                 190

Lys Arg Ser Thr Ile Lys Thr Asp Leu Asp Asp Asp Ile Thr Ala Leu
        195                 200                 205

Lys Gln Arg Ser Phe Val Asp Ala Glu Val Asp Val Thr Arg Asp Ile
    210                 215                 220

Val Ser Arg Glu Arg Val Trp Arg Thr Arg Thr Thr Ile Leu Gln Ser
225                 230                 235                 240

Thr Gly Lys Asn Phe Ser Lys Asn Ile Phe Ala Ile Leu Gln Ser Val
                245                 250                 255

Lys Ala Arg Glu Glu Gly Arg Ala Pro Glu Gln Arg Pro Ala Pro Asn
            260                 265                 270

Ala Ala Pro Val Asp Pro Thr Leu Arg Thr Lys Gln Pro Ile Pro Ala
        275                 280                 285

Ala Tyr Asn Arg Tyr Asp Gln Glu Arg Phe Lys Gly Lys Glu Glu Thr
    290                 295                 300

Glu Gly Phe Lys Ile Asp Thr Met Gly Thr Tyr His Gly Met Thr Leu
305                 310                 315                 320

Lys Ser Val Thr Glu Gly Ala Ser Ala Arg Lys Thr Gln Thr Pro Ala
```

325 330 335

Ala Gln Pro Val Pro Arg Pro Val Ser Gln Ala Arg Pro Pro Pro Asn
340 345 350

Gln Lys Lys Gly Ser Arg Thr Pro Ile Ile Ile Ile Pro Ala Ala Thr
355 360 365

Thr Ser Leu Ile Thr Met Leu Asn Ala Lys Asp Leu Leu Gln Asp Leu
370 375 380

Lys Phe Val Pro Ser Asp Glu Lys Lys Lys Gln Gly Cys Gln Arg Glu
385 390 395 400

Asn Glu Thr Leu Ile Gln Arg Arg Lys Asp Gln Met Gln Pro Gly Gly
405 410 415

Thr Ala Ile Ser Val Thr Val Pro Tyr Arg Val Val Asp Gln Pro Leu
420 425 430

Lys Leu Met Pro Gln Asp Trp Asp Arg Val Val Ala Val Phe Val Gln
435 440 445

Gly Pro Ala Trp Gln Phe Lys Gly Trp Pro Trp Leu Leu Pro Asp Gly
450 455 460

Ser Pro Val Asp Ile Phe Ala Lys Ile Lys Ala Phe His Leu Lys Tyr
465 470 475 480

Asp Glu Val Arg Leu Asp Pro Asn Val Gln Lys Trp Asp Val Thr Val
485 490 495

Leu Glu Leu Ser Tyr His Lys Arg His Leu Asp Arg Pro Val Phe Leu
500 505 510

Arg Phe Trp Glu Thr Leu Asp Arg Tyr Met Val Lys His Lys Ser His
515 520 525

Leu Arg Phe
530

What Is Claimed Is:

1. An isolated nucleic acid comprising a polynucleotide that is at least 90% identical to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 29.

2. An isolated nucleic acid comprising a polynucleotide that is at least 95% identical to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 29.

3. An isolated nucleic acid comprising a polynucleotide that is at least 98% identical to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 29.

4. An isolated nucleic acid comprising RNA corresponding to any of the DNA sequences or fragments of claims 1, 2 or 3.

5. An isolated nucleic acid comprising a DNA sequence identical to a sequence selected from the group consisting of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27 and 28 and the complements of these.

6. An isolated nucleic acid comprising RNA corresponding to the DNA sequence of Claim 5.

7. An isolated nucleic acid comprising at least the polypeptide coding region of a human gene, said human gene containing a DNA sequence according to Claim 1.

8. An isolated nucleic acid comprising at least the polypeptide coding region of a human gene which contains the DNA sequence of Claim 5.

9. The isolated nucleic acid of claim 8 which expresses a human protein when in a suitable expression system.

10. A vector comprising the DNA sequence of claim 1.

11. A vector comprising the DNA sequence of claim 3.

12. A vector comprising the DNA sequence of claim 5.

13. A vector comprising the DNA sequence of claim 9.

14. A polypeptide coded for by the DNA sequence of claim 7 and active fragments, derivatives and functional analogs thereof.

15. A polypeptide coded for by the DNA sequence of claim 8 and active fragments, derivatives and functional analogs thereof.

16. A polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, and 29.

17. A genetically engineered cell having inserted into the genome thereof the DNA of Claim 7.

18. A process for producing cells for expressing a polypeptide using genetically engineered cells of claim 17.

19. An isolated DNA sequence comprising a fragment of a DNA of claim 5, wherein said fragment comprises at least 15 sequential bases of said sequence.

20. An isolated DNA sequence comprising a fragment of DNA of claim 5, wherein said fragment comprises at least 30 sequential bases of said sequence.

21. An isolated DNA sequence comprising a fragment of DNA of claim 5, wherein said fragment comprises at least 50 sequential bases of said sequence.

22. An isolated DNA sequence comprising a fragment of DNA of claim 5, wherein said fragment comprises at least 80 sequential bases of said sequence.

23. A method of detecting genes within the human genome comprising contacting a sample of said genome with an isolated DNA selected from the group consisting of the DNAs of claims 19, 20, 21, and 22.

24. A monoclonal antibody against a polypeptide selected from the group consisting of the polypeptides of claims 14 and 15.

Figure 1

```
GAATTCGGCA CGAGGTCGCG GCGGCGAAGG AGGAGGAGGA AGAGGGCGAG GCGACAAGAG      60
AAGAAGGAGG CAGGCGCGGC GGCAGCGGCG GCGCCCCGAG CCGGCGGAGG CGAGGGGGGG     120
GAAGATGGCG GACGTGCTTA GCGTCCTGCG ACAGTACAAC ATCCAGAAGA AGGAGATTGT     180
GGTGAAGGGA GACGAAGTGA TCTTCGGGGA GTTCTCCTGG CCCAAGAATG TGAAGACCAA     240
CTATGTTGTT TGGGGGACTG GAAAGGAAGG CCAACCCAGA GAGTACTACA CATTGGATTC     300
CATTTTATTT CTACTTAATA ACGTGCACCT TTCTCATCCT GTTTATGTCC GACGTGCAGC     360
TACTGAAAAT ATTCCTGTGG TTAGAAGACC TGATCGAAAA GATCTACTTG GATATCTCAA     420
TGGTGAAGCG TCAACATCGG CAAGTATAGA CAGAAGCGCT CCCTTAGAAA TAGGTCTTCA     480
GCGATCTACT CAAGTCAAAC GAGCTGCAGA TGAAGTTTTA GCAGAAGCAA AGAAACCACG     540
AATTGAGGAT GAAGAGTGTG TGCGCCTTGA TAAAGAGAGA TTGGCTGCCC GTTTGGAGGG     600
TCACAAAGAA GGGATTGTAC AGACTGAACA GATTAGGTCT TTGTCTGAAG CTATGTCAGT     660
GGAAAAAATT GCTGCAATCA AAGCCAAAAT TATGGCTAAG AAAAGATCTA CTATCAAGAC     720
TGATCTAGAT GATGACATAA CTGCCCTTAA ACAGAGGAGT TTTGTGGATG CTGAGGTAGA     780
TGTGACCCGA GATATTGTCA GCAGAGAGAG AGTATGGAGG ACACGAACAA CTATCTTACA     840
AAGCACAGGA AAGAATTTTT CCAAGAACAT TTTTGCAATT CTTCAATCTG TAAAAGCCAG     900
AGAAGAAGGG CGTGCACCTG AACAGCGACC TGCCCCAAAT GCAGCACCTG TGGATCCCAC     960
TTTGCGCACC AAACAGCCTA TCCCAGCTGC CTATAACAGA TACGATCAGG AAAGATTCAA    1020
AGGAAAAGAA GAAACGGAAG GCTTCAAAAT TGACACTATG GGAACCTACC ATGGTATGAC    1080
ACTGAAATCT GTAACGGAGG GTGCATCTGC CCGGAAGACT CAGACTCCTG CAGCCCAGCC    1140
AGTACCAAGA CCAGTTTCTC AAGCAAGACC TCCCCCAAAT CAGAAGAAAG GATCTCGAAC    1200
ACCCATTATC ATAATTCCTG CAGCTACCAC CTCTTTAATA ACCATGCTTA ATGCAAAAGA    1260
CCTTCTACAG GACCTGAAAT TTGTCCCATC AGATGAAAAG AAGAAACAAG GTTGTCAACG    1320
AGAAAATGAA ACTCTAATAC AAAGAAGAAA AGACCAGATG CAACCAGGGG GCACTGCAAT    1380
TAGTGTTACA GTACCTTATA GAGTAGTAGA CCAGCCCCTT AAACTTATGC CTCAAGACTG    1440
GGACCGCGTT GTAGCCGTTT TTGTGCAGGG TCCTGCATGG CAGTTCAAAG GTTGGCCATG    1500
GCTTTTGCCT GATGGATCAC CAGTTGATAT ATTTGCTAAA ATTAAAGCCT TCCATCTGAA    1560
GTATGATGAA GTTCGTCTGG ATCCAAATGT TCAGAAATGG GATGTAACAG TATTAGAACT    1620
CAGCTATCAC AAACGTCATT TGGATAGACC AGTGTTCTTA CGGTTTTGGG AAACATTGGA    1680
CAGGTACATG GTAAAGCATA AATCGCACTT GAGATTCTGA ATTATTTGGC TCCTCCATTT    1740
CTGGAAATTG AGACTCAAGC TTTATGAATT TATCAAGAAC TTAAAAATGA AGAAGGTCAC    1800
AGATTGATCT TTTATAAGAC CTTATTTGAT GCTTTGTGCT TCAAGGAGAT GATACCTGTC    1860
ATCCATATAA GCAAACTTTT TGGCTTACAA CTATTTTTTT AATATTAGCC TTCTAGTCTG    1920
TAATGGAAAT TGTATATTTT GATAGAAGTT TTTTCTCCAT TGGTTAAATT AGCATTACTT    1980
AAAATTTGTT TCTTTAGAAA ATAAATGCAG GTTATAAATG TGTGTATATT TAGAGATTAT    2040
AAGGCTCTCT GAGCCATCTT CTGATTTTTC ATTGCTCTAT AATTCTTTTT ACTGAAAATA    2100
CTATGTTATG AATGGTATTA AATTTTAGTC TCTGGAACAT CCAAAACCAA GCAAAGGGAT    2160
GTGACTATTT TGAATGAATC AGAATGTCAA CTTGTATGTA CACTATATCT ACACTTACTC    2220
ATTATTTAAA AAGAATAATG AAAAATCTAG ATCAATTCTT CAATTTGATT GAACTGTTCA    2280
GCCTTTTCAA GATTTCTTTA TTTACAAATG ATTACATTTA AATGAATGTA CATTCTTCTC    2340
ACTGACTTTG GTGATTTTGA AACCTAGAAT GATGTGTTTC TATCTGTAAT ATCTTTCCAT    2400
TTGAAAAAAA TCTCAAAACA CAGATTAAAA CCACAAAAAA AAAAAAAAAA AAA           2453
```

Figure 2

MADVLSVLRQYNIQKKEIVVKGDEVIFGEFSWPKNVKTNYVVWGTGKEGQPREYYTLDSI
LFLLNNVHLSHPVYV*RRAATENIPVVRRPDRKD*LLGYLNGEASTSASIDRSAPLEIGL
QRSTQVKRAADEVLAEAKKPRIEDEECVRLDKERLAARLEGHKEGIVQTEQIRSLSEAMS
VEKIAAIKAKIMAKKRSTIKTDLDDDITALKQRSFVDAEVDVTRDIVSRERVWRTRTTIL
QSTGKNFSKNIFAILQSVKAREEGRAPEQRPAPNAAPVDPTLRTKQPIPAAYNRYDQERF
KGKEETEGFKIDTMGTYHGMTLKSVTEGASARKTQTPAAQPVPRPVSQARPPPNQKKGSR
TPIIIIPAATTSLITMLNAKDLLQDLKFVPSDEK*KKQGCQRENETLIQRRKD*QMQPGG
TAISVTVPYRVVDQPLKLMPQDWDRVVAVFVQGPAWQFKGWPWLLPDGSPVDIFAKIKAF
HLKYDEVRLDPNVQKWDVTVLELSYHKRHLDRPVFLRFWETLDRYMVKHKSHLRF

Figure 3

Human RNA Master Blot.

	1	2	3	4	5	6	7	8
A	whole brain	amygdala	caudate nucleus	cerebellum	cerebral cortex	frontal lobe	hippocampus	medulla oblongata
B	occipital pole	putamen	substantia nigra	temporal lobe	thalamus	subthalamic nucleus	spinal cord	
C	heart	aorta	skeletal muscle	colon	bladder	uterus	prostate	stomach
D	testis	ovary	pancreas	pituitary gland	adrenal gland	thyroid gland	salivary gland	mammary gland
E	kidney	liver	small intestine	spleen	thymus	peripheral leukocyte	lymph node	bone marrow
F	appendix	lung	trachea	placenta				
G	fetal brain	fetal heart	fetal kidney	fetal liver	fetal spleen	fetal thymus	fetal lung	
H	yeast total RNA	yeast tRNA	E. coli rRNA	E. coli DNA	Poly r(A)	human C0t DNA	human DNA	human DNA

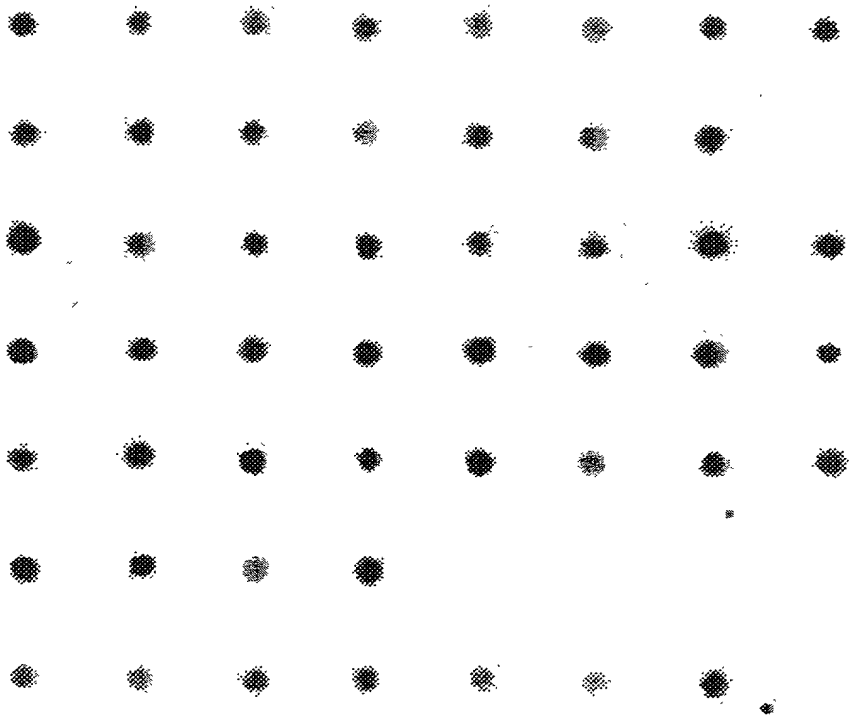


Figure 4(A)

DSC64: Tissue Distribution

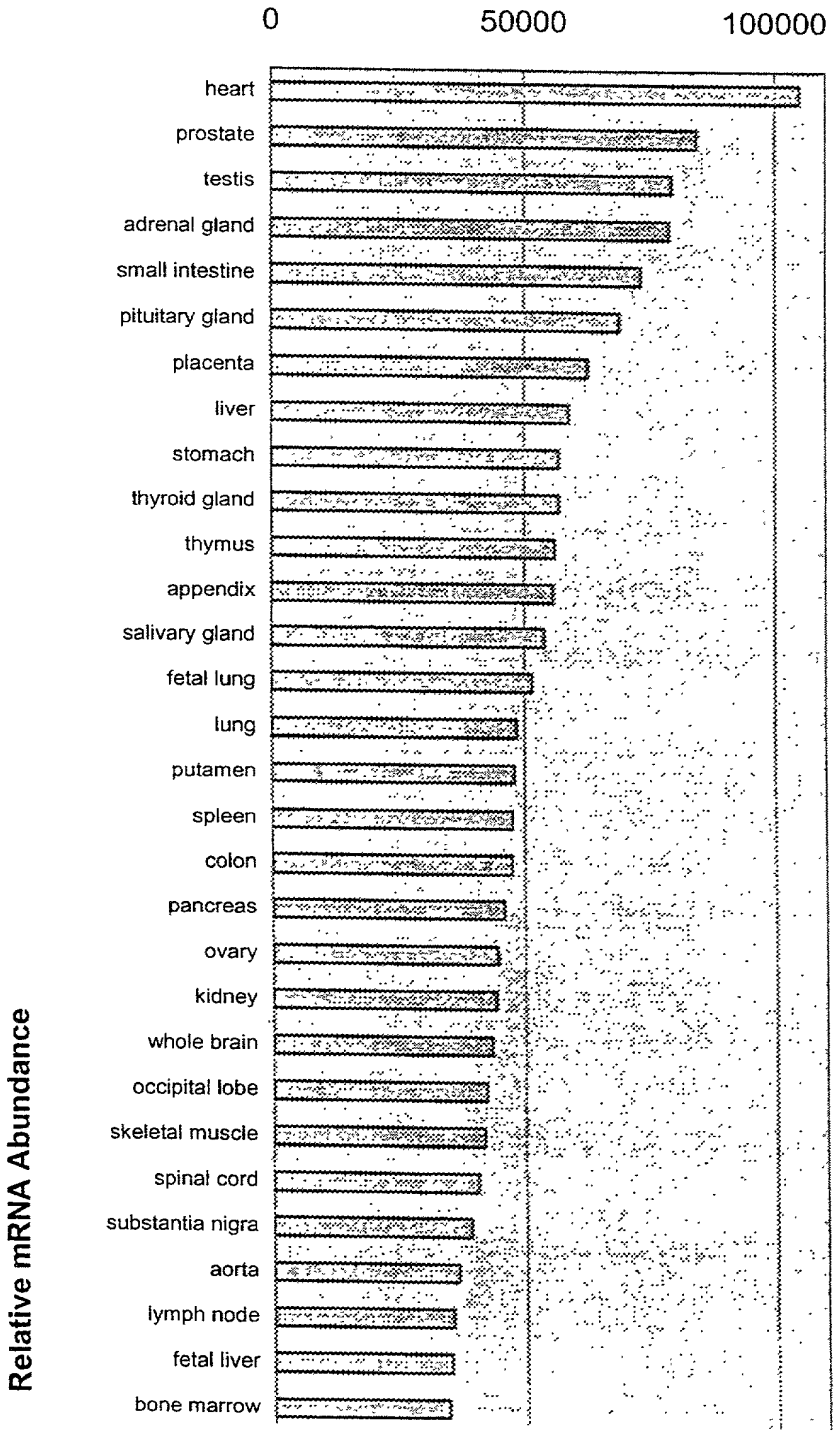
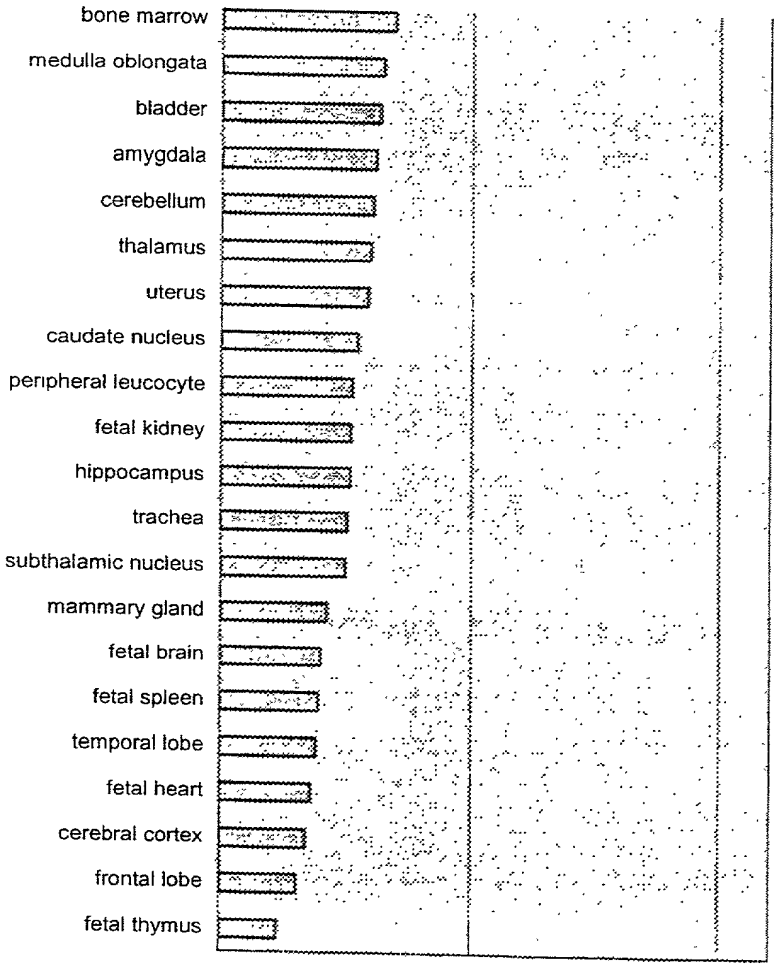


Figure 4(B)


bone marrow
medulla oblongata
bladder
amygdala
cerebellum
thalamus
uterus
caudate nucleus
peripheral leucocyte
fetal kidney
hippocampus
trachea
subthalamic nucleus
mammary gland
fetal brain
fetal spleen
temporal lobe
fetal heart
cerebral cortex
frontal lobe
fetal thymus

COMBINED DECLARATION FOR PATENT APPLICATION AND POWER OF ATTORNEY (Includes Reference to PCT International Applications)	ATTORNEY'S DOCKET NUMBER 640100-430 Customer No.: 27162

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

Human Mesenchymal DNAs and Expression Products

the specification of which (check only one item below):

☐ is attached hereto.

☒ was filed as United States application
Serial No. 09/937,974
on October 1, 2001
and was amended on (if applicable)

☒ was filed as PCT international application
Number PCT/US00/08751
on 31 March 2000
and was amended under PCT Article 19
on (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the patentability of this application in accordance with Title 37, Code of Federal Regulations §1.56(a).

I hereby claim foreign priority benefits under Title 35 United States Code §119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed:

PRIOR FOREIGN/PCT APPLICATION(S) AND ANY PRIORITY CLAIMS UNDER 35 U.S.C. 119:

COUNTRY (if PCT indicate PCT)	APPLICATION NUMBER	DATE OF FILING (day, month, year)	PRIORITY CLAIMED UNDER 35 USC 119	
			☐ YES	☐ NO
			☐ YES	☐ NO
			☐ YES	☐ NO
			☐ YES	☐ NO
			☐ YES	☐ NO

COMBINED DECLARATION FOR PATENT APPLICATION AND POWER OF ATTORNEY
(Includes Reference to PCT International Applications)

ATTORNEY'S DOCKET NUMBER
640100-430
Customer No.: 27162

I hereby claim the benefit under Title 35, United States Code, §120 or § 119 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

PRIOR U.S. APPLICATION(S) OR PCT INTERNATIONAL APPLICATIONS DESIGNATING THE U.S. FOR BENEFIT UNDER 35 U.S.C. 120 or U.S.C. 119:

U.S. APPLICATIONS		STATUS (Check one)		
U.S. APPLICATION NO.	U.S. FILING DATE	PATENTED	PENDING	ABANDONED
60/127,418	1 April 1999		☐	☐
60/148,800	13 August 1999			
PCT APPLICATIONS DESIGNATING THE U.S.			☐	☐
PCT APPLICATION NO.	PCT FILING DATE	U.S. SERIAL NUMBERS ASSIGNED (if any)		
PCT/US00/08751	31 March 2000		☒	☐

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith. (List name and registration number) John N. Bain (Reg. No. 18,651); John G. Gilfillan III (Reg. No. 22,746); Elliot M. Olstein (Reg. No. 24,025); Raymond J. Lillie (Reg. No. 31,778); William Squire (Reg. No. 25,378); Alan Grant (Reg. No. 33,389), Francis C. Hand (Reg. No. 22,280) and Glennon Troublefield (Reg. No.39,050)

Send Correspondence to: Alan J. Grant, Esq.
Carella, Byrne, Bain, Gilfillan, Cecchi, Stewart & Olstein
6 Becker Farm Road, Roseland, New Jersey 07068

Direct Telephone Calls to:
(name and telephone number)
(973) 994-1700

201	FULL NAME OF INVENTOR	FAMILY NAME Van den Bos	FIRST GIVEN NAME Christian	SECOND GIVEN NAME
	RESIDENCE & CITIZENSHIP	CITY Baltimore	STATE OR FOREIGN COUNTRY Maryland MD	COUNTRY OF CITIZENSHIP Germany
	POST OFFICE ADDRESS	POST OFFICE ADDRESS 2214 East Fairmont Avenue	CITY Baltimore	STATE & ZIP CODE/COUNTRY Maryland 21231
202	FULL NAME OF INVENTOR	FAMILY NAME Mbalaviele	FIRST GIVEN NAME Gabriel	SECOND GIVEN NAME
	RESIDENCE & CITIZENSHIP	CITY O'Fallon	STATE OR FOREIGN COUNTRY Missouri MO	COUNTRY OF CITIZENSHIP Republic of Congo
	POST OFFICE ADDRESS	POST OFFICE ADDRESS 8 Chastain Court	CITY O'Fallon	STATE & ZIP CODE/COUNTRY Missouri, 63366
203	FULL NAME OF INVENTOR	FAMILY NAME	FIRST GIVEN NAME	SECOND GIVEN NAME
	RESIDENCE & CITIZENSHIP	CITY	STATE OR FOREIGN COUNTRY	COUNTRY OF CITIZENSHIP
	POST OFFICE ADDRESS	POST OFFICE ADDRESS	CITY	STATE & ZIP CODE/COUNTRY

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application of any patent issuing thereon.

SIGNATURE OF INVENTOR 201	SIGNATURE OF INVENTOR 202	SIGNATURE OF INVENTOR 203
DATE 11/27/01	DATE 12-04-01	DATE